

July 2019

## The Teacher Evaluation Conundrum: Examining the Perceptions of Special Education Teachers

Gordon Brobbey

*University of South Florida*, brobbeyg@yahoo.com

Follow this and additional works at: <https://digitalcommons.usf.edu/etd>



Part of the [Special Education and Teaching Commons](#)

---

### Scholar Commons Citation

Brobbey, Gordon, "The Teacher Evaluation Conundrum: Examining the Perceptions of Special Education Teachers" (2019). *USF Tampa Graduate Theses and Dissertations*.

<https://digitalcommons.usf.edu/etd/8339>

This Dissertation is brought to you for free and open access by the USF Graduate Theses and Dissertations at Digital Commons @ University of South Florida. It has been accepted for inclusion in USF Tampa Graduate Theses and Dissertations by an authorized administrator of Digital Commons @ University of South Florida. For more information, please contact [digitalcommons@usf.edu](mailto:digitalcommons@usf.edu).

The Teacher Evaluation Conundrum: Examining the Perceptions of Special Education Teachers

by

Gordon Brobbey

A dissertation submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy in Curriculum and Instruction  
with a concentration in Special Education  
Department of Teaching and Learning  
College of Education  
University of South Florida

Major Professor: Ann Cranston-Gingras, Ph.D.  
Robert Dedrick, Ph.D  
Brenda Walker, Ph.D.  
Vonzell Agosto, Ph.D

Date of Approval:  
June 20, 2019

Keywords: accountability, evaluation systems, teacher effectiveness, students with disabilities.

Copyright © 2019, Gordon Brobbey

## **DEDICATION**

This dissertation is dedicated to the most meaningful people in my life. I dedicate it to my wife Elizabeth Brobbey, my daughters Melanie and Mirabel Brobbey, my parents Isaac Oteng and Vida Oteng Asuamah, my siblings Ama, Kofi, Yaa, and Abenaa Yeboaa, and finally my outstanding major professor, Dr. Ann Cranston-Gingras. You are my heroes. Thank you for believing in me and inspiring me.

## **ACKNOWLEDGEMENTS**

The most fortunate and impactful coincidence, that will remain etched in my memory, was the assignment of Dr. Ann Cranston-Gingras as my initial advisor at the start of my doctoral journey in August 2015. I had never met Dr. C.G, as you are affectionately called, even though I did my masters in the department a few years prior. After my first few meetings with you as my advisor, I knew instinctively that you would be the right person to guide me through the tumultuous years that lay ahead in doctoral life. You would prove to be the most approachable, motivating, understanding, and encouraging professor I had ever met. I was amazed by your unapparelled humility and meekness in spite of your impressive academic and administrative credentials. It was an easy decision to make when the time finally came to choose a major professor. You became my academic ‘mother,’ mentor, and guide. Thank you, Dr. C.G. for the gentle, timely nudges, motivation, pieces of advice, and the professional exposure and opportunities that have shaped my scholarship up to this point. I will forever remain indebted to you.

Dr. Walker, I am not sure you recollect this, but you sowed the seed of this adventure in one of the courses I took with you during my Master’s studies. You inspired me to pursue this dream with all confidence. Thank you for time as well as the different forms of practical guidance that watered and natured the seed you sowed into a fruitful one. To Dr. Dedrick, thank you for accepting to be on my committee as the methodologist. You were reassuring and graceful every time I spoke with you. Dr. Agosto, thank you for all the constructive dialogues you had with me. It was almost predictable that every time we had an intellectual conversation, I

would become more knowledgeable than I was before the conversation. I am grateful for your genuine interest in my growth both as a scholar and a professional. To Dr. Elizabeth Doone and Dr. Jeannie Kleinhammer-Tramill (Dr. K.T.), thank you for the timely pieces of advice and all the opportunities for growth you provided me.

I have reached this milestone because of the incredible support I received while in K-12 as a special education teacher. I wish to express my appreciation to Mr. Steve van Gorden, a former principal at Zephyrhills High School in Pasco County Schools. Steve, despite all the perceived odds and impediments at the time, you discerned what was not obvious to many and gave me opportunity to prove myself. To Tammy Hochstetler of Zephyrhills High School, you were such a great mentor. Michael Pellegrino of R.B. Steward Middle School, you have a great heart. Thanks for your generosity. Bridget White, thanks for being a reliable colleague. You have no idea how much inspiration and encouragement you have given me all these years.

I was also fortunate to be part of the 2015 close-knit cohort of special education doctoral students at the University of South Florida. I appreciate your friendship and mutual encouragement.

I would further like to express appreciation to the participants who took time to respond to the survey. The rich perspectives you shared with me will shape by teaching, scholarship, and community engagement in the years ahead.

To my family, I could not have attained this feat without your iron-clad love and support. You were the refuge for me as I battled the stormy moments of doctoral studies. Throughout the high and low moments, you stuck by me and urged me on. To ‘Grandma’ Suletha Hosey, my wife Elizabeth Brobbey, and my daughters, Mirabel Brobbey and Melanie Brobbey, I say “Me da mo ase” (i.e. thank you in Twi).

## TABLE OF CONTENTS

TABLE OF CONTENTS.....	ii
LIST OF TABLES.....	iv
LIST OF FIGURES.....	v
ABSTRACT.....	vi
CHAPTER ONE: INTRODUCTION.....	1
Statement of the Problem.....	5
The Challenges Surrounding Special Education Teaching and Evaluation.....	6
Paucity of Research on Special Education Teacher Perceptions.....	10
Attrition in Special Education.....	11
Purpose of the Study and Research Questions.....	13
Significance of the Study.....	14
Definition of Terms.....	16
Conceptual Framework.....	18
Delimitations.....	22
CHAPTER TWO: LITERATURE REVIEW.....	23
Chapter Overview.....	23
Comprehensive Literature Review.....	23
Inclusionary and Exclusionary Criteria.....	23
Literature Search.....	25
Findings.....	27
Issue Relating Defining Effective Teaching in Special Education.....	27
Features of Effective Special Education Teaching.....	29
Effective Special Education Teaching and Existing Systems.....	31
Current Evaluation Methods and Associated Challenges.....	34
Special Education Teacher Responsibilities and Roles.....	39
Proposed Evaluation Approaches.....	44
Status of Research on Stakeholder Perceptions.....	51
Discussion.....	60
Summary of Chapter.....	64

CHAPTER THREE: METHOD .....	65
Introduction.....	65
Methods and Procedures .....	67
Methods.....	67
Instrument .....	68
Population and Sampling .....	70
Human Subjects Approval .....	74
Procedures .....	74
Ethical Implications .....	77
Data Analysis .....	78
Summary .....	81
 CHAPTER FOUR: RESULTS .....	 82
Overview.....	82
Results.....	84
Summary of Results .....	99
 CHAPTER FIVE: DISCUSSION.....	 102
Overview.....	102
Findings.....	103
Understanding of evaluation system and components.....	108
Perceptions about Roles, Responsibilities, and Contexts .....	110
Implications for Practice and Policy .....	118
Recommendations for Future Research .....	123
Limitations of Study .....	124
Summary of Chapter Five.....	125
Conclusion .....	126
 REFERENCES .....	 128
 APPENDICES .....	 138
Appendix 1: Survey Instrument .....	139
Appendix 2: Informed Consent.....	147
Appendix 3: Initial Invitation to Participants.....	150
Appendix 4: First Reminder Email .....	151
Appendix 5: Second Reminder Email.....	152
Appendix 6: University IRB Approval.....	153

## LIST OF TABLES

Table 1: Key Terms for Literature Search .....	25
Table 2: School District Enrollment Demographics.....	72
Table 3: Perceptions of Clarity .....	85
Table 4: Participant's Overall Perceptions of Clarity.....	85
Table 5: Practicality: Perceptions about Evaluator.....	86
Table 6: Practicality: Overall Perceptions about Evaluator .....	87
Table 7: Practicality: Perceptions about Feedback.....	88
Table 8: Practicality: Overall Perceptions about Feedback .....	89
Table 9: Practicality: Perceptions about Student Growth.....	90
Table 10: Practicality: Overall Perceptions about Student Growth .....	91
Table 11: Practicality: Perceptions about Professional Development .....	91
Table 12: Practicality: Overall Perceptions about Professional Development .....	92
Table 13: Cost Perceptions .....	93
Table 14: Overall Cost Perceptions .....	94
Table 15: T-test: Group Statistics of Clarity, Practicality, and Cost by Gender.....	95
Table 16: One-Way ANOVA of Clarity, Practicality, and Cost Perceptions by Years of Experience.....	96
Table 17: One-Way ANOVA of Clarity, Practicality, and Cost Perceptions by Job Categories ..	98



## **LIST OF FIGURES**

Figure 1: Factors Influencing Teacher Perceptions of Policy.....	19
Figure 2: Conceptual Framework of Special Education Teacher Perceptions about Evaluation Systems .....	21
Figure 3: Literature Search Process .....	26

## **ABSTRACT**

For more than a decade now, state and local education agencies have adopted high-stakes teacher evaluation systems in response to federal accountability mandates and the recognition of the critical role of teacher effectiveness in student achievement. Teacher evaluation systems have revealed relevant details regarding teacher effectiveness, especially in general education settings. However, no systems of teacher evaluation have been developed and validated for special education teachers and little is known about how special education teachers have fared under the accountability microscope. The study was informed by three significant issues pertaining to special education teacher evaluation: (1) the challenges surrounding evaluation of special education teacher effectiveness, (2) limited research documenting special education teacher views, perceptions, voices, and experiences pertaining to how they have fared under existing evaluations systems, and (3) the persistent problem of special education teacher attrition within the context of high-stake teacher evaluations. Special education teachers are typically evaluated with measures designed for use with general education teachers, and researchers have pointed out that those measures do not address their unique professional development and pedagogical needs.

The purpose of the study was to explore how special education teachers perceived existing teacher evaluation systems that are used to evaluate their effectiveness. The study examined how special education teachers view the evaluation systems with regard to the significance and ability of the systems to distinguish the multiple roles, responsibilities, and contexts within which they work.

Using quantitative methods, the study employed a census survey to solicit the perceptions of secondary level special education teachers from one school district in southwest Florida. A total of 357 participants were asked to respond to an electronic survey about the perceptions of the teacher evaluation system used in their district. The number of respondents included in the final data analysis was 96, representing a 26.8% response rate.

An instrument was adopted and revised for data collection. The reliability of the revised scales as measured by Cronbach's alphas ranged from .70 to .89. Findings from the study reveal that participants did not have a deep understanding of the evaluation system used to assess their performance, especially as it relates to how all components are combined to generate summative performance scores. The findings also reveal that participants did not view the evaluation framework as practical to address their roles, responsibilities, and professional development needs. Although participants did not view the evaluation system as practical to their needs and jobs, they had a slightly favorable view of their evaluators' ability to assess them in a fair and unbiased manner. As a result, participants strongly recommended for a separate evaluation system for special education teachers. Participants in the study further reported a high emotional cost of stress and anxiety associated with the evaluation process.

The findings have implications for federal, state, and local education agencies and policy makers. Recommendation for future research are also discussed.

## **CHAPTER ONE: INTRODUCTION**

Efforts to document teacher behavior and productivity in the classroom are probably as old as the teaching profession itself. According to Hazi and Arredondo Rucinski (2009), teacher supervision and evaluation have long been part of educational research as a result of their promise to improve teachers and their classroom practice. The history of teacher evaluation indicates that the focus of evaluations has shifted over time. For example, teacher evaluations in the 1940s and 1950s emphasized teacher traits such as trustworthiness, warmth, and enthusiasm as benchmarks of teacher effectiveness (Danielson & McGreal, 2000). The 1960s and 1970s ushered an era of evaluations that focused on teachers' ability to enhance the acquisition of basic skills. That era also saw increased research in clinical supervision where researchers began developing observation instruments that provided "accurate depictions of what was occurring in the classrooms" (Danielson & McGreal, 2000, p. 13; Watts, 2016). By the 1980s and 1990s, however, the focus on teacher evaluation had shifted from teacher-centered traits to student-centered outcomes. The demands of a changing job market that required students to graduate with complex skills such as critical thinking, problem-solving, and collaborative learning were cited as the rationale for the shift (Danielson & McGreal, 2000).

For more than a decade now, however, "teacher evaluation has become a serious business" (Grissom & Youngs, 2016, p. 1) in view of the tremendous emphasis on accountability and the role of teacher effectiveness in student achievement (Johnson & Semmelroth, 2014a).

The passage of the No Child Left Behind Act (NCLB) in 2001 initiated an era of the federal government's focus on accountability, teacher effectiveness, and the teacher's role in students' academic achievement (Bacon, 2015; Jones & Brownell, 2014). Two provisions, specifically, in NCLB elevated the need for comprehensive teacher evaluation systems: Adequate Yearly Progress (AYP) and 'highly qualified' teacher requirements. The AYP requirement emphasized the need for all students, including students with disabilities, to be 100% proficient in reading, as demonstrated through standardized assessments, by the 2013-2014 school year (Tandy, Whitford, & Hirth, 2016). The attainment of such an ambitious goal was contingent on the quality of teachers in the classroom (Hazi & Arredondo Rucinski, 2009; Tandy, Whitford, & Hirth, 2016), hence the requirement that all teachers have the requisite educational qualifications, subject matter expertise, and meet state licensure expectations in order to fulfill the 'highly qualified' teacher mandate.

Other federal mandates such as the 2006 Teacher Incentive Fund grants and the 2009 Race to the Top competitions further raised the stakes on teacher quality and accountability (Brownell & Jones, 2015; Glowacki & Hackmann, 2016; Quigney, 2010). Through the Teacher Incentive Fund grants, for instance, states and local education agencies that made the commitment to link teacher performance to incentives received funding from the federal government. According to the US Department of Education (2018), the purpose of the fund was to support the use of performance-based pay and other human capital strategies, with the view to increasing students' access to quality teachers. Although the program was revised in 2017 and renamed 'Teacher and School Leader Incentive Program Grant' competition, it still retains the

original focus of promoting performance-based compensation for teachers, principals, and other school leaders (US Department of Education, 2018). The Race to the Top program, with its emphasis on teacher evaluation, compensation, and retention policies (Quigney, 2010), also rewarded states for developing teacher evaluation systems that incorporated student achievement measures in evaluating teacher effectiveness (Brownell & Jones, 2015; Woolf, 2013). It specifically required applying states to revise their teacher and administrator evaluation policies to include considerations for student growth data (Darling-Hammond, 2013; Glowacki & Hackmann, 2016).

These policies illustrate the shift to highlighting teacher effectiveness and quality as the primary contributors to student success. Granted, there are several contributors to student achievement including family dynamics, poverty, in-school interventions, and teacher characteristics. In terms of in-school interventions, however, researchers have concluded that teacher effectiveness has the greatest impact on student achievement (Council for Exceptional Children, 2012; Jones & Brownell, 2014; National Council for Accreditation of Teacher Education, 2010). The available research thus suggest that teacher performance and effectiveness are the vitally important predictors of students' academic achievement (Gordon, Kane, & Staiger, 2006; Holheide, Goe, Croft, & Reschly, 2010).

NCLB also elevated concerns about the poor performance of US students compared with their international counterparts, thus the need to raise performance standards to enable US students to become competitive on the global academic scene (Bacon, 2015; Woolf, 2015). In response to this concern, together with the critical role of the teacher in student achievement, policy makers and philanthropic organizations, such as the Melinda and Bill Gates Foundation, have advocated for and promoted teacher evaluation systems and value-added models (VAM) as

the primary accountability tools to ameliorate the perceived deficits in teacher performance and to boost students' academic attainment (Darling-Hammond, 2013). As a result of these initiatives, states and school districts have, since 2009, instituted evaluation systems that employ multiple approaches to measure teacher performance and effectiveness. According to Brownell and Jones (2015), 38 states and the District of Columbia have revised their evaluation systems to require the inclusion of student achievement measures in the assessment of teacher effectiveness due to federal policies, and all 50 states require the use of observations in determining teacher effectiveness.

The increased focus on teacher accountability and subsequent adoption and use of high-stakes evaluation systems by states and school districts have not been without controversy and concerns. One of the major concerns relates to the use of value-added models (VAMs), a summative performance evaluation, used to estimate student growth on standardized assessments as a conduit to assess teacher effectiveness by way of a teacher's contribution to the measured growth. Using VAMs to assess teacher performance has been controversial on a number of grounds. For one thing, teachers lack understanding of models because they are based on complex statistical formulas. Teachers also disagree with the extent to which standardized assessments accurately capture the type of learning that is seen as important (Cohen & Goldhaber, 2016). In addition, there is debate regarding the ability of VAMs to estimate the contributions of each teacher, as well as variables outside the control of the school or teacher, to student achievement (Cohen & Goldhaber, 2016; Lawson, 2014; Moran, 2017).

Issues surrounding teacher evaluations become even more complex when discussed within the context of special education. Researchers have pointed out that current evaluation systems, which were designed to evaluate general education teachers, do not address the peculiar

pedagogical approaches of special education teachers or the instructional needs and outcomes of students with disabilities (Johnson, Crawford, Moyland, & Ford, 2016; Watts, 2016). Concerns have also been raised about using VAMs to estimate special education teacher effect on student achievement (Jones & Brownell, 2014; Johnson et al., 2016). Additional concerns have been raised about the multiplicity of roles and responsibilities assigned to special education teachers and their implications for assessing special educator effectiveness and quality (Glowacki & Hackmann, 2016; Johnson et al., 2016; Jones, 2016; Tandy, Whitford, & Hirth, 2016; Watts, 2016). Given the implications of high-stakes evaluations for special education teacher performance, compensation, and job prospects, it is vital to elicit the views of special educators in terms of how they have fared under existing evaluation systems.

### **Statement of the Problem**

This dissertation research is premised on three significant problems in relation to special education teacher evaluation. They include: (1) the challenges surrounding evaluation of special education teacher effectiveness, (2) the paucity of research documenting special education teacher views, perceptions, voices, and experiences in terms of how they have fared under existing evaluations systems, and (3) the persistent problem of special education teacher attrition within the context of high-stake teacher evaluations. The urgent need to research in these problematic issues is necessitated by the fact that “special education teachers have generally been omitted from large-scale studies, and only a handful of empirical studies have examined issues raised in the evaluation of such teachers” (Jones, 2016, p. 63). The problems underpinning this dissertation are discussed in detail in the sections that follow.



## **The Challenges Surrounding Special Education Teaching and Evaluation**

In response to accountability demands and scrutiny into teacher quality, there has been a proliferation of teacher evaluation systems in recent years to evaluate teacher performance. One sub-group of teachers has, however, not been given due attention in these efforts. According to Jones and Brownell (2014), teacher evaluation systems have revealed vital insights concerning teacher effectiveness in the general education setting, especially in mathematics and language arts. In contrast, the same cannot be said about special education teachers. No systems of evaluation have been developed for use with special education teachers or the varied settings in which they teach (Woolf, 2015). In fact, current mechanisms for evaluating general education teachers “have not been validated for use with special education teachers, and their designs do not adequately address the characteristics that make up the field” (Johnson & Semmelroth, 2014, p. 67). School districts typically rely on approaches that have only been validated for use with general education teachers.

Although there is near unanimity in the special education literature regarding the inappropriateness of using system that are validated for use in general education settings for special educators, little effort has been expended towards developing systems that recognize the uniqueness of special education teachers’ responsibilities and roles (Holdheide, 2015). In a report commissioned by the National Comprehensive Center for Teacher Quality, Holdheide, Goe, Croft, and Reschly (2010) surveyed more than 1,100 state and district level directors of special education and conducted interviews with school level administrators regarding challenges in evaluating special education teachers’ performance. Holdheide et al. (2015) were also interested in establishing the status of district and state policies in terms of promising special education evaluation practices. Holdheide et al. (2015) found that more than half of the school districts

surveyed developed their own evaluation systems and another 33.3 percent used or adopted systems recommended by the state. A noteworthy finding was that 71.9 percent of respondents indicated that contractual agreements barred districts from modifying or differentiating evaluations instruments for special education teachers. Ironically, nearly half of respondents (49.9%) were of the view that special education and general education teachers should not be evaluated with the same system. These findings clearly illustrate some of the concerns that special educators and researchers have expressed about teacher evaluation systems that disregard the need to differentiate for the roles and responsibilities of special educators.

One of the reasons attributed to the seeming neglect of special education teachers in the design and validation of evaluation measures is the challenge of evaluating special education teachers due to the complex nature of their roles and responsibilities (Holheide et al, 2010; Jones, 2016; Jones & Brownell, 2014; Johnson & Semmelroth, 2014). Most states have adopted either the Danielson Framework for Teaching (FFT) or the Marzano Teacher Evaluation Model. According to Brownell and Jones (2014), these two observation protocols are similar in content and structure but their validity has not been assessed for use with special education teachers although states and districts continue to use them for that purpose.

Researchers have raised a number of challenges in relation to special education teacher evaluation (Brownell & Jones, 2015, Johnson et al., 2016, Johnson & Semmelroth, 2014, Watts, 2016). These challenges include: different groups of students in varying contexts, nature of instruction in special education, and competing demands on the special education teacher (Johnson & Semmelroth, 2014).

***Varying Contexts:*** According to Johnson and Semmelroth (2014), students with special needs constitute approximately 12% of the total population of students in the United States.

Within this relatively small population, there is wide variation in terms of instructional contexts within which this group is served. To meet students' needs, the special education teacher may have to collaborate with the general education teacher in the general education setting.

Alternatively, the special education teacher may have to pull students from the general education setting into resource rooms to be able to provide specialized instruction. Some special education teachers also provide instruction in self-contained settings or in extended resources room in order to facilitate learning for students with severe needs. Special educators may also work in consulting roles as support facilitators, providing support to both the general education teacher and students. Of interest is the fact that special education teachers sometimes have to combine a number of these varying roles, especially in smaller and resource-deprived school districts. The heterogeneity of these roles makes it difficult to evaluate teacher quality in special education, thus requiring a flexible approach than what is currently provided by existing evaluation models (Council for Exceptional Children, 2012; Jones & Brownell, 2014; Sledge & Pazey, 2013).

The evident complexities in the special educator's roles within varying contexts make it problematic to attribute the gains or lack of achievement to the special education teacher, given that students with special needs may also receive instruction from the general education teacher. Hence, Jones and Brownell (2014) contended, "when we observe with either setting, we may be seeing an incomplete picture of instruction that student receive in a given subject. And, a teacher's effectiveness in providing instruction for students is inevitably going to depend, at least in part, on the quality of instruction that they received in the other setting" (p. 113).

***Nature of Special Education Instruction:*** Providing students with specialized, individualized instruction is a basic premise in special education teaching. This fundamental characteristic of special education instruction is rooted in the mandates of the Individuals with

Disabilities Act (IDEA, 2014) that, among other things, requires instruction to be delivered to students in harmony with the dictates of the Individual Education Plan (IEP) to suit the needs of each student (CEC, 2012; Johnson, 2016). This requirement demands that special education teachers be well versed in multiple instructional strategies and practices suitable for different disability categories. As a result of the varying needs of students, individualized instruction at times, may also require individualized goals and outcomes (Johnson & Semmelroth, 2014; Woolf, 2015). In addition to the individualistic nature of instruction, special education pedagogy favors direct, explicit instruction in view of the needs of students with learning needs. On the other hand, most of the existing evaluations systems such as the FFT and Marzano's Framework for Teaching were designed in the constructivist paradigm where students take charge of their own learning. As a result of the complex nature of individual needs among students with disabilities, Johnson and Semmelroth argue against "defining student achievement through one universal measure, or even through a set of accepted predetermined measures" (p. 73) as set out in teacher evaluation systems. Therefore, any teacher evaluation models that purport to measure special education teacher competence must consider all the challenges associated with teaching students with a wide array of needs in diverse settings (Johnson & Semmelroth, 2014).

***Competing Demands on the Special Educator:*** A final challenge associated with evaluating special education teachers is the multiple demands on their time. Apart from the instructional time spent on teaching in varying settings, special education teachers provide other necessary services such as mentoring and coaching to students with disabilities (Council for Exceptional Children, 2012). Additional responsibilities for special educators include case management, instructional remediation, planning, writing, and coordinating IEP meetings, and collaboration with other stakeholders such as agencies, general education teachers, speech and

language therapists, parents, and paraprofessionals (Jones & Brownell, 2014). These competing, yet necessary demands, take a toll on how much time the special educator is able to spend on instruction. According to one estimate, the special educator spends only 16% of class time on instruction (Vannest & Hagan-Burke, 2009). In addition, Johnson and Semmelroth (2014) pointed out that caseload requirements have increased over the past twenty years despite research to suggest that larger caseloads and instructional numbers negatively impact student achievement, at least in math and reading.

### **Paucity of Research on Special Education Teacher Perceptions**

Research conducted in general education has revealed important insights regarding teacher effectiveness in terms of how teacher evaluation systems can be used to make high-stakes decisions about teachers (Jones, 2016; Jones & Brownell, 2014). On the contrary, there is a dearth of studies measuring special education teacher perceptions of teacher evaluation systems, in part due to the omission of special education teachers from large-scale studies that have examined issues relating to the evaluation of such educators (Jones, 2016). There have been a handful of studies (e.g., Glowacki & Hackman, 2016; Lawson, 2015; Lawson & Knollman, 2017; Rasul, 2018; Widener, 2011) that examined administrators' perceptions of the special education teacher evaluation process. On the other hand, studies examining special educators' perceptions, perspectives, opinions, and views are almost non-existent. Only three studies (i.e., Quartico, 2016; Doer, 2012; Jiang, Spote, & Luppescu, 2015) were found to have examined special education teacher perceptions. Two of the three studies (i.e., Quartico, 2016; Doerr, 2012) were dissertation studies. The third study (Jiang et al., 2015) was conducted with all teachers, including some special education teachers as participants. Jiang et al.'s (2015) study examined teacher perspectives on evaluation reforms in Chicago Public Schools. Jiang et al.

(2015) drew from 32 interviews from a random sample of teachers and 2 years of survey data from more than 12,000 per year over a period of two years to measure their perceptions about teacher evaluation systems. Although the study participants were drawn from all groups of teachers, Jiang et al. found that special education and high school teachers were significantly not in favor of using student growth measures to evaluate teachers. They also found out that teachers who received school-wide value-added scores had more negative views about being held accountable for the work of other teachers and students they were not associated with. The finding regarding how special education teachers perceived the evaluation system is significant, considering the unique challenge of attributing student growth to a given teacher. Special education instruction is typically nested between general and special education, especially for students in high incidence disability categories (Brownell & Jones, 2015). Teachers' perceptions about the evaluation process and instruments may impact their beliefs in relation to the types of pedagogical decisions they make (Milanowski, 2017).

Acknowledging the importance of teacher inputs and views in the design and implementation of evaluation systems, the CEC (2012) recommended that special educators must be included in the development, implementation, evaluation of the teacher evaluation process. The need to understand teacher perceptions is amplified when the implications of evaluation systems for teacher tenure, remuneration, promotion, certification, and retention are considered (Darling-Hammond, 2013).

### **Attrition in Special Education**

Teacher attrition rates, a significant contributor to the ever-increasing problem of teacher shortage, have seen a rising trend over the years and continue to do so (Kelchtermans, 2017). The trend has been historically pervasive in special education (Boe, 2014). According to Sutchter,

Darling-Hammond, and Carver-Thomas (2016), 48 states reported shortages in special education. Although shortages were also reported in Mathematics and Science, special education sub-categories comprised more than half of all severe shortage areas (Sutcher, Darling-Hammond, & Carver-Thomas, 2016). In Florida, the trend is not different. The Florida Department of Education (FLDOE, 2018), in reporting teacher shortage areas by certification, indicated that special education ranked fourth (4<sup>th</sup>) in relation to certification areas representing the greatest need among teachers statewide. For the 2014-2015 school year, 8.32% of special education courses were taught by uncertified special education teachers. The field is also beset with job dissatisfaction and teachers who are certified through alternative means (Johnson & Semmelroth, 2014).

There is research to show that the nature and quality of training for special educators, together with years of experience positively impact student achievement (Sledge & Pazez, 2013). According to Jones and Brownell (2014), although there are few studies examining the effect of teacher characteristics on achievement in special education, the available evidence suggest that preparation and years of experience have some effect on special education teachers compared with little or no preparation. Yet, some researchers (e.g., Gordon, Kane, & Staiger, 2006) have suggested that the bottom 25% of teachers who are found ineffective, based on evaluation systems and value-added models, should be replaced. Johnson and Semmelroth argued against this stance by maintaining “we cannot fire special education teachers as the sole or even primary means of improving the quality of the teaching system – who will replace them?” (p. 73). What is needed, in place of firing, are evaluation systems that will provide meaningful feedback to support special education teachers in improving their practice and addressing their pedagogical needs and professional development aspirations.

## **Purpose of the Study and Research Questions**

The purpose of the study was to explore special education teachers' perceptions about teacher evaluation systems that are used to assess their performance and effectiveness. The study examined how special education teachers view teacher evaluation systems in relation to their significance and ability to distinguish the multiple roles and contexts within which they work. The study further examined whether special education teachers see current evaluation models as fitting enough to measure their competence and address their pedagogical and professional development needs. Given that teacher evaluation policies have taken center-stage across the nation, with implications for teacher remuneration, job retention, and student achievement (Darling-Hammond, 2013), it is pertinent to consider the views of the individuals (teachers) upon whom the success or failure of policies depend (Jiang et al., 2015). The following research questions guided the study:

- How do special education teachers rate their understanding of the teacher evaluation system used in their district?
- What are the perceptions of special education teachers about teacher evaluation systems with regard to their multiple roles, responsibilities, and contexts?
- What are the perceptions of special education teachers about evaluations systems with regard to the ability of these systems to address professional development, practice, and pedagogical needs?
- Do special education teachers' perceptions of teacher evaluation systems vary by:
  - a) job categories
  - b) gender
  - c) years of experience.



## **Significance of the Study**

Teacher evaluation systems have been adopted by school districts across the country in the past few years in response to federal policies on accountability and teacher effectiveness. For special educators, Holtheide (2015) identified three different strategies states and districts have adopted for evaluating them: (1) using one model for all teachers, both general and special educators, (2) using a differentiated model that is supplemented with special considerations and examples for special education teachers, and (3) adopting separate models for different categories of teachers. What is not known, however, is how special education teachers have fared under these different systems in comparison to their general education counterparts. There is also limited research about how special education teachers have responded to various evaluation systems. The limited research base emanates from the fact that special education teachers have generally not been included in large-scale empirical studies that have examined issues related to teacher evaluation. As a result, much is not known about whether existing systems being used “lead to valid, reliable information on special educator effectiveness” (Jones, 2016, p. 63). According to Jones (2016), there is “long commentary surrounding the challenges of evaluating special education teachers but short on empirical evidence that can guide policymaking...we need evidence related to how the FFT and other protocols, as currently implemented in school, function with special educators” (Jones, 2016, p. 72). Empirical studies are needed to understand how special education teachers have responded to these systems and the extent to which they have adjusted their pedagogy and professional development considerations to align with what is considered effective teaching on evaluation rubrics (Holdheide, 2015).

In addition, there is a lack of consensus on how special education teachers should be evaluated (Johnson & Semmelroth, 2014) owing to the complexities associated with the nature of

pedagogical approaches in special education, varying instructional contexts, and the competing demands from the special educator. Some researchers (e.g. Brownell & Jones, 2015; CEC, 2012; Holdheide et al., 2010, Jones & Semmelroth, 2014) have called for a replacement or modification of existing systems and rubrics to better account for special educators. Others (i.e., Sledge & Pazez, 2013) have recommended the provision of adequate training for assessors and principals to help them develop the requisite expertise to accurately assess special educator quality and to provide useful feedback. There are challenges associated with measuring special educator effectiveness and quality. The dilemma over how special education teachers must be evaluated is complicated by the complexities of responsibilities, the lack of research on what constitutes effective instruction in special education, and deficiencies in the teacher current evaluation methods.

Researchers and policy makers have recognized the need to develop systems that give attention to the peculiar nature of the special education field. In response to this need, some researchers have created and validated at least two observation protocols specifically for special educators. First, Johnson and Semmelroth (2014) developed the Recognizing Effective Special Education Teachers (RESET) protocol for the state of Idaho. According to Johnson and Semmelroth (2014), the protocol was designed “as a possible alternative to measure special education teacher effectiveness (p. 71). Although this tool fills a critical void, it has been criticized for its sole focus on the instructional component of special education teaching and learning while ignoring the social dimensions (Barnes, Cipriano, Flynn, Rivers, & Xu 2018). The second one is the Recognizing Excellence in Learning and Teaching (RELATE) observation tool (Barnes, Cipriano, Flynn, Rivers, & Xu 2018) designed to make up for the shortfalls of RESET. According to Barnes et al. (2018), the goal of the RELATE tool is to “provide a more careful

examination of both instructional and social processes...in the special education classroom” (p. 2). One limitation, however, is that RELATE is designed only for use in self-contained settings, thus limiting its usefulness for the varied instructional settings within which special educators teach. It is also worthy to point out that none of these tools have been adopted for use nationwide. There is, therefore, an urgent need for research to ascertain how special education teachers have fared under evaluation systems that are not developed or validated for them. It is surprising that, to date, no approaches have been validated for use with special education educators (Brownwell & Jones, 2015; Johnson & Semmelroth, 2014; Jones & Brownell, 2014; Sledge & Pazez, 2013) across the country.

Researchers have also documented the importance of teacher perceptions in the successful implementation of policy initiatives (Jiang, Spote, & Luppescu, 2015). Teachers are the gatekeepers and final decision makers on classroom practice and they resist change and reform initiatives if they are not aligned with their perceptions and beliefs (Doyle & Ponder, 1977; Muncey & McQuillan, 1996). Teacher evaluation has been promoted as the accountability mechanism through which to ensure the promotion of teacher quality needed for student achievement and success. Given the increased number of students with special needs, about 12% of the entire US student population, to be specific, who are being educated in today’s classrooms across the country (Glowacki, 2013), there is an urgent need to research teacher perceptions and recommendations that can be used to inform the development of evaluation systems.

### **Definition of Terms**

The following terms are used in the study and operationalized as follows:

Special education teacher      A teacher who is certified by the State of Florida in the areas of Special Education K-12 and whose primary

responsibility is to teach students who are diagnosed with a disability classified under IDEA and have an Individual Education Plans (IEP), regardless of their disability category.

Teacher evaluation

A formal process that a district or school employs to review and evaluate teachers' performance and effectiveness in the classroom (Editorial Projects in Education Research Center, 2015)

Teacher Effectiveness

The term teacher effectiveness is difficult to define, especially so within the context of special education. What characterizes effectiveness in general education cannot be juxtaposed to special education. For the purpose of this dissertation, Jones and Brownell (2014) definition of effective special education teaching will be used. Teacher effectiveness in special education is characterized by a teacher's ability to deliver instruction that is (a) explicit, (b) cohesive, (c) intensive, (d) engaging, (e) responsive to student needs, and (f) focused on essential concepts, strategies, and skills.

Perception

As defined by Keenan (2018), perception refers to how people organize, identify, and interpret information gathered through the senses. It allows people to make meaning out of what they experience in the world. Perception also deals with ideas created in the mind based on context, personal experience, and expectations.

## Conceptual Framework

This study of special education teachers' perceptions about teacher evaluation systems drew from Fullan (2001) and Doyle and Ponder (1977) conceptualization of teacher perceptions on change. Policies and reforms, in general, including teacher accountability measures and teacher evaluations systems, are intended to change and improve teacher performance and quality. The success or failure of new policies or reforms are contingent on how teachers perceive those initiatives. Doyle and Ponder (1977) posited that teachers' perceptions and reactions to change proposals and policies that seek to alter their classroom practices is influenced by the extent to which they perceive the policies as practical. Practicality in this context was defined as "an expression of teacher perceptions of the potential consequences of attempting to implement a change proposal in the classroom" (Doyle & Ponder, 1977, p. 6). Doyle and Ponder (1977) argued that policy recommendations that teachers perceive as practical are the ones they are likely to incorporate in classroom procedures and practices. Teachers' determination of practicality is based on the three criteria of *congruence*, *instrumentality*, and *cost*. *Instrumentality* means that a proposed change, policy, or reform must describe procedures in terms that depict classroom practices. *Congruence* refers to the extent to which proposed change, policy or reform is consistent with teachers' perceptions of their own prevailing conditions. Finally, *cost* is explained as the ease with which a proposed change, policy, or reform can be implemented and the potential return teachers can anticipate for adoption the change.

Fullan (2001) also posited that teachers' reaction to policy change and implementation is influenced by four factors. These factors include: need, clarity, complexity, and practicality and quality of proposed change. *Need* refers to whether the proposed policy or reform addresses what teachers perceive to be a priority. *Clarity* refers to how clear a policy's intentions and

means are to teachers (Fullan, 2001; Jiang et al., 2015). *Complexity*, according to Fullan (2001) refers to the “difficulty and extent of change required of the individual responsible for implementation” (p. 78). By *practicality and quality*, Fullan (2001) meant the need to give attention to quality, availability of resources, materials, and time needed to ensure successful adoption and implementation of change. The factors influencing the implementation change, policy, and reform are illustrated in Figure 1.

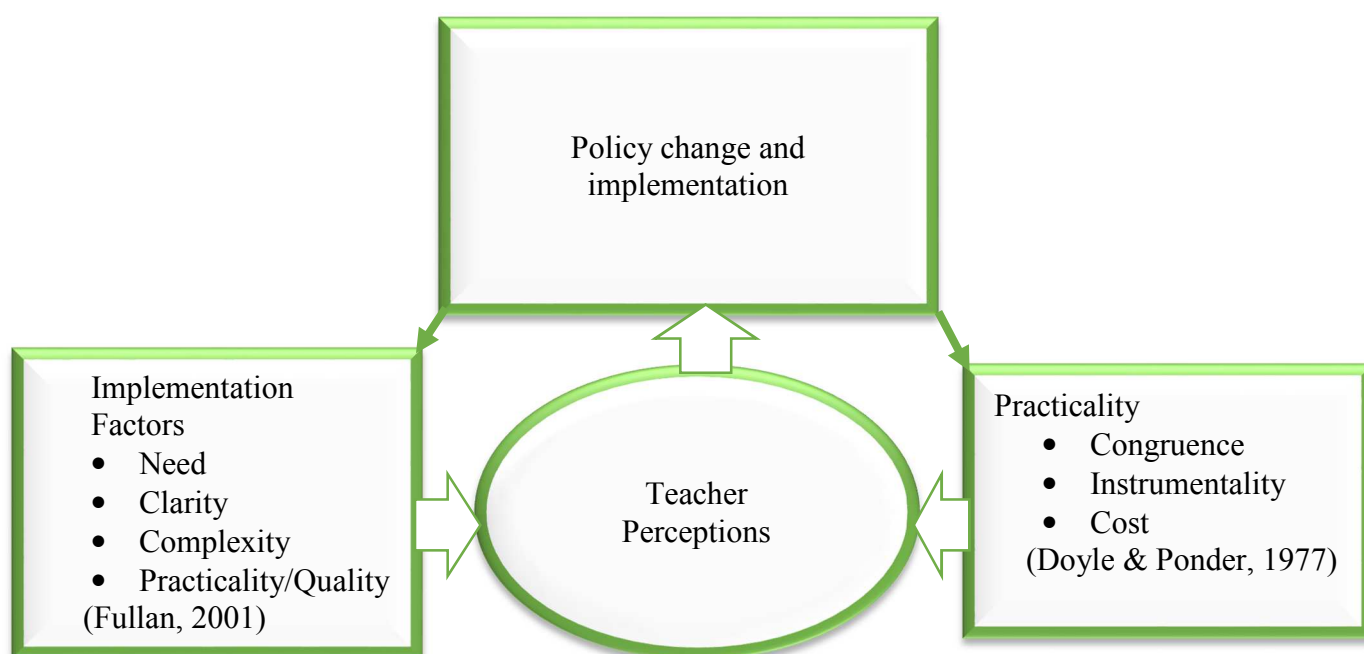


Figure 1: *Factors Influencing Teacher Perceptions of Policy*

Doyle and Ponder’s (1977) and Fullan’s (2001) theory of change has been used to conceptualize at least two studies in connection with teacher evaluation in the broader context of general education (e.g., Jiang et al., 2015; Pizmony-Levy & Woolsey, 2017). It has not been used in connection with special education teacher evaluation, however. Jiang et al. (2015) drew theoretically from Fullan, and Doyle and Ponder to conceptualize their study of teacher perceptions regarding Chicago Public School’s Recognizing Educators Advancing Chicago

Students (REACH) teacher evaluation system. Jiang et al. (2015) utilized the critical factors of implementation identified by Fullan (2001) due to their relation to the extent to which teachers adjust their practice, beliefs, and response to change. In harmony with the instrument used, the Jiang et al. (2015) adopted two out of the four factors of implementation identified by Fullan (2001). The factors included clarity and practicality. Jiang et al. (2015) further used the three dimensions of instrumentality, congruence, and cost to define *practicality* and to frame their study. The researcher utilized this theoretical framework to conceptualize this study. It seems likely that special education teachers will not adjust their practices and pedagogical decisions if they do not perceive teacher evaluation rubrics as practical to assess their performance and job descriptions.

Pizmony-Levy and Woolsey (2017) studied the attitudes of teachers towards high-stakes teacher accountability policies and the New Jersey teacher evaluation reform called the Teacher Effectiveness and Accountability for Children of New Jersey Act (TEACHNJ) using the policy implementation framework. The study reported that the majority of teachers opposed the evaluation system and that teachers' perceptions of the policy were mediated by the politics of the major advocates of the policy, perceptions of implementation efforts, and beliefs in the possible outcomes of the policy.

For the purpose of this study, the theories of implementation as advanced by Fullan (2001) and Doyle and Ponder (1977) are combined to frame and interpret the findings of how special education teachers perceive existing teacher evaluations systems. The conceptualization is as shown in Figure 2.

Figure 2 is based on the conceptualization that success or failure of teacher evaluation reforms, particularly as used for special education teachers, is related to special educator

perceptions. The implementation of these reforms would be influenced by practicality, clarity, and cost perceptions special education teachers have towards these initiatives. Practicality, as used here includes the dimensions of instrumentality and congruence, based on the items on the instrument adopted (Jiang et al., 2015) and the purpose of this study.

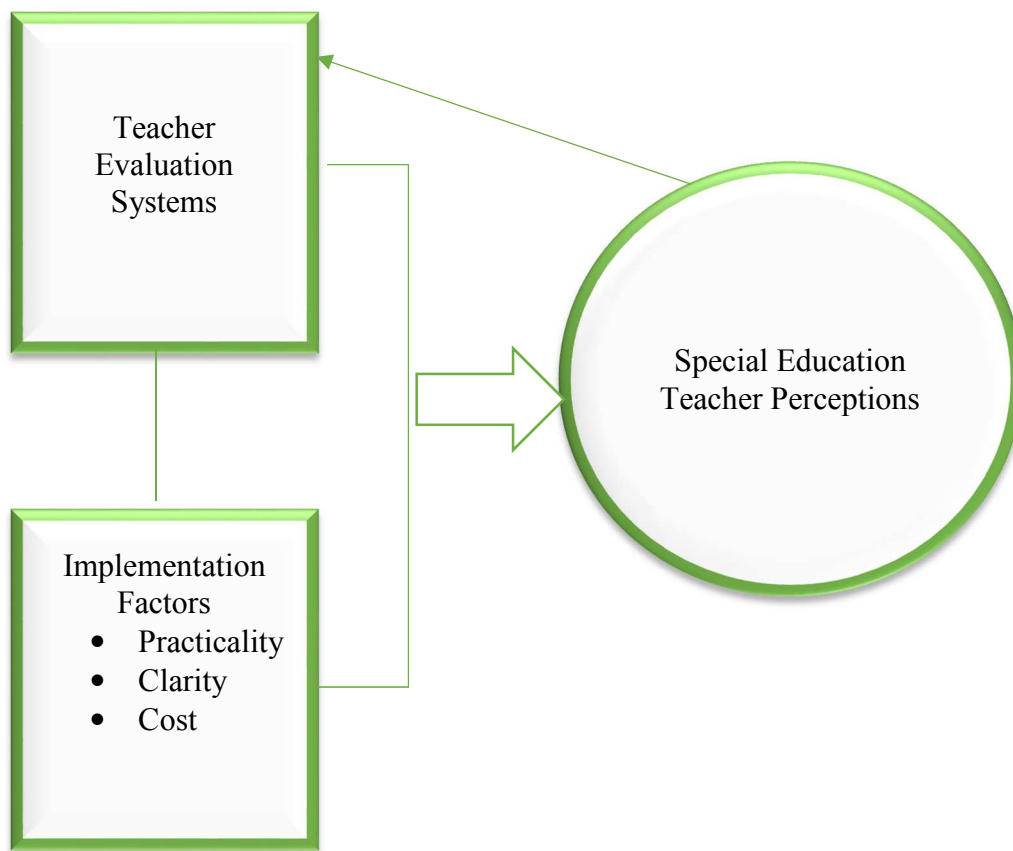


Figure 2: *Conceptual Framework of Special Education Teacher Perceptions about Evaluation Systems*

This study of special educators' perceptions of teacher evaluation systems contributes to the nascent research on how they should be evaluated and inform ongoing policy discussions regarding teacher evaluation reforms in view of recent changes authorized by the 2015 passage



of the Every Student Succeeds Act (EESA). It also contributes to knowledge base in special education and fill the gaps in the literature on teacher evaluation for special education teachers.

### **Delimitations**

The delimitation of the study included methodology, sample size, study location, and focus. The study utilized survey methods to elicit special education teachers' perceptions. The selection of participants was not based on the statistical considerations of power, effect size, and alpha level, as is typically the case with basic, non-experimental quantitative study. Instead, a census survey approach, where all participants in a small and easily-identifiable sample are surveyed was used. The final number of participants was 96.

The study was also bound by the conceptual framework adopted. The conceptual framework was guided by the instrument used, with the recognition that other studies using the theories of implementation advanced Doyle and Ponder (1977) and Fullan (2011) may arrive at different conceptualizations based on the instruments used.

The study was conducted in one school district in southwest Florida and focused on the perceptions of selected special education teachers in the middle and high schools in the selected district. The study did not consider issues relating to diversity, contextual factors at school site, administrator perceptions, or other school-related issues.

## **CHAPTER TWO: LITERATURE REVIEW**

### **Chapter Overview**

This chapter presents a review of literature on the issue of special education teacher evaluation. A thorough review of the literature revealed the status of teacher evaluation efforts in special education and also frames the methods informing this study. Boote and Beile (2005) underscored the importance of a comprehensive literature review, which serves as the foundation to conducting a study in any field. According to Boote and Beile (2005), “a comprehensive, thorough, sophisticated literature review is a precondition for doing substantive, thorough, sophisticated research” (p. 3). This chapter, therefore, is an effort to understand what has been studied and written on the topic of special education teacher evaluation, including the strengths and weaknesses of methods used in the literature. The review on special education teacher evaluation commences with a discussion of the inclusionary and exclusionary criteria followed by the search strategies used to retrieve the literature. Next, a discussion of the findings from the review and their relationship to the purpose of the study are presented. A summary of the chapter is also provided.

### **Comprehensive Literature Review**

#### **Inclusionary and Exclusionary Criteria**

Prior to commencing the search for literature, a set of criteria were designed to sort through the voluminous amount of special education literature. These criteria were set on the

premise of suitability and quality. In their framework for determining the standards and criteria for literature reviews, Boote and Beile (2005) developed a five-category framework: coverage, synthesis, methodology, significance, and rhetoric. The first category, coverage, deals with the extent to which the researcher justified the criteria for inclusion and exclusion. To justify studies selected for the review, a set of four inclusionary/exclusionary criteria used by Guarino, Santibanez, and Daley (2006) in their review of relevant literature on teacher retention and recruitment was adopted. The criteria involved: (1) relevance, (2) scholarship, (3) empirical nature, and (4) quality.

In line with the above criteria, studies for inclusion in the review were deemed relevant if they gave evidence of ability to provide context/background for the research questions, focused on in-service special education teachers (Guartico, 2015), and were published from 2010 to 2018. The choice of 2010 to 2018 enables the selection of studies that reflect the most recent and up-to-date trends within the special education evaluation policy context.

For a determination of the scholarliness, studies from peer-reviewed journals, scholarly articles from reputable and professional special education organizations and policy institutions were included. In addition, recent dissertations and book chapters that provided empirical evidence relevant to the topic were included.

Studies were considered empirical if they utilized qualitative, quantitative, or mixed methodological approaches and offered empirical findings and conclusions. For the final criteria of quality, studies that presented rigorous lines of argument, findings, and supported conclusions in harmony with the current special education literature were considered. The literature base in

special education, especially in connection with teacher evaluation, is limited and sparse. Therefore, conceptual and descriptive articles, as well as policy briefs from prominent and nationally-recognized policy organizations were selected for inclusion. Studies that dealt with the evaluation of pre-service teachers, general education teachers, or school supporting staff and service providers like school psychologists, guidance counsellors, speech-language pathologists, and occupational therapists were excluded from this review.

### Literature Search

The search for studies that met the inclusionary criteria began with the identification of the key concept, key words, phrases, and search terms that would yield maximum and quality retrieval of literature. The key concept in this study, and which was used for the search, was *special education teachers*. This key concept was paired with key words and phrases such as *teacher evaluation, evaluating special education teachers, evaluating co-teachers, evaluation, accountability, perceptions, perspectives, and viewpoints, and opinions*. Various permutations of the search terms were made using ‘*and*’ and ‘*or*’ to arrive at the final set of studies selected in line with the inclusionary criteria. The table below shows the permutations

Table 1: *Key Terms for Literature Search*

Key Concept	AND	Key Phrases	OR	Key Words
Special education teachers		Teacher evaluation		Perceptions
		Evaluating special education teachers		Perspectives
		Evaluating co-teachers		Viewpoints
		Evaluation		Opinions
		Accountability		

The search began with the entry of the search terms and phrases and the different combinations into the following electronic databases: ERIC, Google Scholar, Education Source, PsycInfo, and SAGE Research Methods. In addition, the “Cited By” feature of Google Scholar and as well as citation digging from the various studies were utilized to retrieve similar studies. The selection process that culminated in the final selection of studies is illustrated in Figure 2.

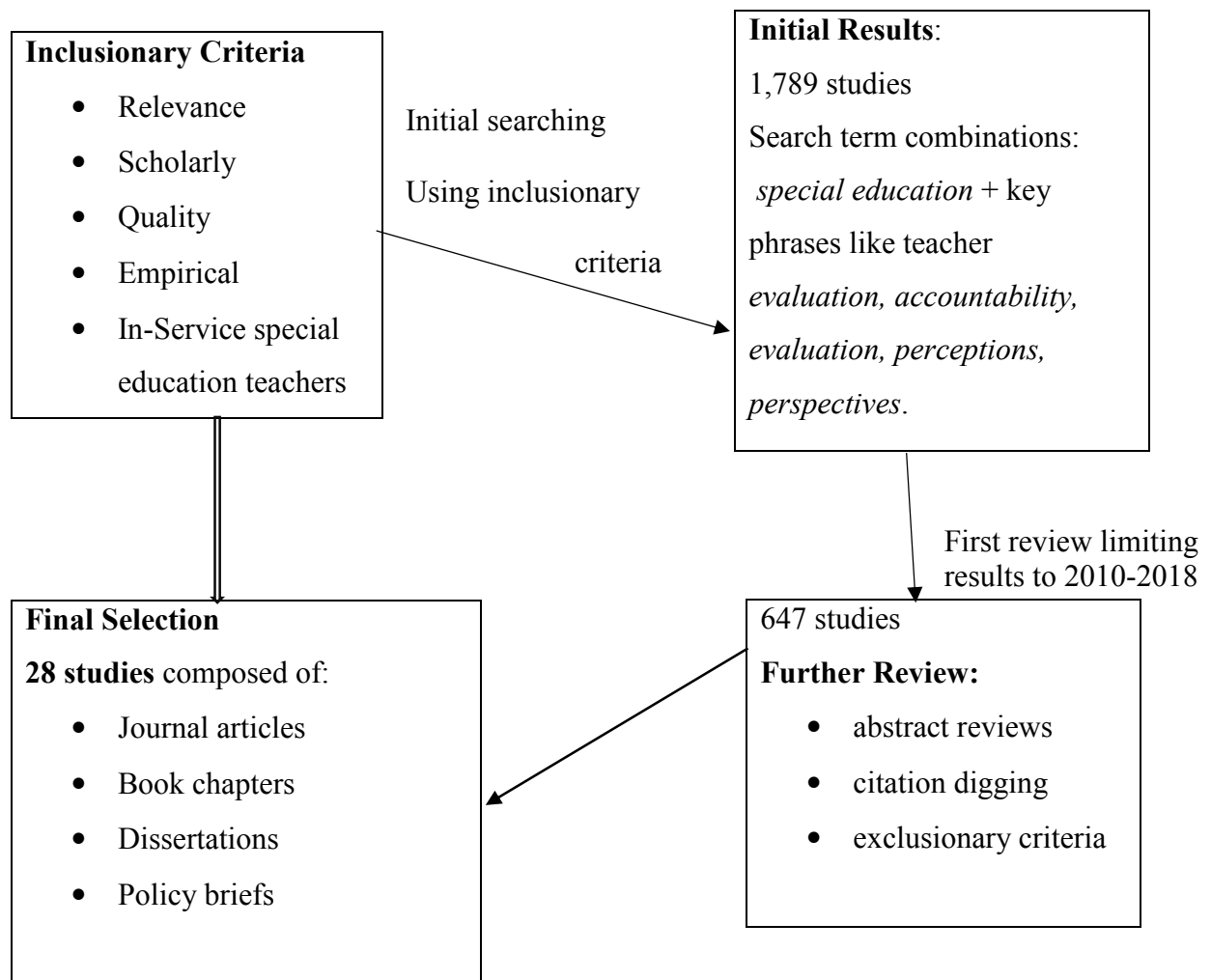


Figure 3: *Literature Search Process*

The initial search, using different permutations of *special education* and *teacher evaluation* yielded 1,789 results. The resulting studies were pruned down to 647 after limiting

the results to studies between 2010 and 2018. After further screenings and reviews in line with the inclusionary/exclusionary criteria and the purpose of the study, 28 studies were finally selected to be included in the literature review.

### **Findings**

Four major themes emerged from the comprehensive review of the literature on special education teacher evaluation. The themes include (a) issues relating to defining what is considered effective teaching in special education, (b) current evaluation methods for special education teachers and associated challenges, (c) proposed evaluation approaches to remedy challenges associated with existing evaluation systems, and (d) status of research on stakeholder perceptions of the evaluation process. In a study on special education teachers' perceptions of the design and implementation of the special education evaluation process, Guartico (2016) reviewed the special education literature from 2000 to 2016 and found similar themes, which were: characteristics of effective teaching in special education settings, current methods of evaluation for special educators, and perceptions regarding current special education systems. Guartico (2016) did not, however, find a theme on proposed measures to evaluate special educators. The four major themes from the review covering the special education literature on evaluating special educator effectiveness from 2010 to 2018 are synthesized and presented.

#### **Issue Relating Defining Effective Teaching in Special Education**

Teacher effectiveness has been at the forefront of the movement to hold teachers accountable for the educational outcomes of their students. If evaluation efforts are to succeed in sifting effective teachers from ineffective ones, there must be clear standards regarding expectations of performance. The need for clear standards of performance becomes even more pertinent in the case of special educators for whom no evaluation measures have been developed,

and yet, are expected to demonstrate specialized expertise (Woolf, 2018). The findings in the special education literature on issues relating to what competencies special education teachers are expected to demonstrate to be considered effective are presented below.

All studies under this theme discussed features of special education teaching within the context of teacher evaluation for the purposes of informing the design or modification of evaluation systems for use with special educators. Some studies had a nationwide focus while others were restricted to specific states. The Council for Exceptional Children (2012), with a focus on nationwide evaluation systems, pointed out a number of measures of special education teacher effectiveness that must be included in designing evaluation measures. Johnson (2015) described the features of effective special education teaching within the framework of developing a localized system of evaluation for the state of Idaho. Jones and Brownell (2014) proposed a definition of effective special education teaching and used the definition as the basis of investigating an existing, and commonly-used evaluation tool across multiple states and school districts. Sledge and Pazey (2013) discussed special education teacher effectiveness within the context of a broader nationwide reform efforts in teacher evaluation and how special education teachers can be included in such reforms.

The CEC published its popular *The Council of Exceptional Children's Position on Special Education Teacher Evaluation* in 2012 to advance the professional body's views on considerations for ensuring equitable treatment and inclusion of special education teachers into existing teacher evaluating systems that had, hitherto, not factored special educators' roles and responsibilities into designing and implementing teacher evaluation models. The CEC (2012) argued that a valid teacher evaluation can serve as the means to judge a special education teacher's knowledge and skills hence the need for states and local education agencies to

consciously integrate precise standards and evidence-based practices that delineate what is expected of special educators. As part of the five-point position, the CEC (2012) stated that teacher evaluations systems must use evidence-based practices that provide multiple indicators of special education teacher effectiveness. The indicators of special education teacher effectiveness, according to the CEC (2012) included: (a) development and implementation of IEPs, (b) development of lesson plans, (c) skills in providing students with disabilities access to the general education classroom and curriculum, (d) classroom environment and management, (e) identification and implementation of the requisite and appropriate instructional strategies, (f) progress monitoring and assessment, (g) collaboration with colleagues and families, and (h) participating in ongoing professional development. The CEC (2012), however, fell short of offering specific guidelines regarding how these evidence-based practices can be integrated into the different evaluation systems and tools adopted by states and districts, and which although not validated for use with special educators, continue to be utilized to assess special education teacher performance and effectiveness.

### **Features of Effective Special Education Teaching.**

Jones and Brownell (2014), on the other hand honed in on a widely-adopted evaluation tool, the Charlotte Danielson's Framework for Teaching (FFT), and focused their study on steps that would be required to validate the tool for use with special education teachers. Jones and Brownell (2014) reviewed the empirical research base in special education to frame their investigation into what constitutes effective teaching in special education. The researchers preceded their definition of effective special education teaching by first distinguishing between the *teaching quality* and *teacher quality*. The authors stated that *teaching quality* is dependent on the context within which instruction occurs. As such, characteristics of students, the school, and



the larger context needs to be taken into consideration. Context is particularly important in special education where research is complex to conduct and assessing teacher quality is even more complex owing to the multiplicity of roles and responsibilities assigned to special educators. In relation to *teacher quality*, that tends to deal with how prospective teachers are trained, Brownell and Jones (2014) noted that the research base on the characteristics of a quality special education teachers is very limited. The limited research base, however, reveals that special education coursework, certification status, degree earned, teacher preparation route, and years of experience appear to be important predictors of special education teacher quality. With the distinction between teaching quality and teacher quality in special education in mind, Brownell and Jones (2014) proposed a definition of special education teaching. According to them, effective special education teaching is defined by six features, which include: (1) explicit, (2) cohesive, (3) intensive, (4) engaging, (5) responsive to student needs, and (6) focused on essential concepts, strategies, and skills.

*Explicit* instruction, as defined by the authors, involves stating a rationale for learning, modelling, giving clear explanations, and practicing with students how to use and apply a concept, skill or strategy with new tasks. Each phase of explicit instruction is marked by high level of interaction. *Cohesive, intensive, and engaging* special education instruction, according to Jones and Brownell (2014) is marked by special education teachers' ability to connect big ideas during instruction, doing so across different instructional sessions, and with intensity for the most at-risk students. Instruction is also marked by seamless management of transitions with less time wasted, focused and repeated practice, and active student cognitive engagement. By *responsive* special education teaching, the authors referred to special education teachers' ability to arrange, modify, and adjust instruction and instructional processes and practices, using student

assessment data, in response to their students' needs. The final characteristic, *focused* instruction, was explained as special education teachers' ability to "provide instruction focused on essential concepts, strategies, and skills" (Jones & Brownell, 2014 p. 116).

### **Effective Special Education Teaching and Existing Systems.**

After proposing their framework for defining effective special education teaching, Jones and Brownell (2014) compared the definition with Danielson's FFT rubric to evaluate its feasibility in assessing special education teacher effectiveness. The authors found that five out of the six defining features of effective special education teaching appeared to align with the FFT's rubrics, but one feature was missing, an observation the researchers found problematic. The missing feature was *explicitness*, and while it appeared to be just one feature, it is the most significant or core feature of special education instruction. Students with disabilities generally benefit from direct instruction, in which the teacher plays a central role in explaining concepts, modeling them, and giving ample opportunities for practice until students can comprehend and apply the concept. The authors argued that the absence of explicitness, which is the hallmark feature of effective teaching in special education, undermined the FFT's validity for use in evaluating special education teacher effectiveness. The authors were of the view that the FFT's language used to describe optimum teacher performance is focused on students' actions that are independent of teacher support, an expectation that may be problematic for students with disabilities who thrive on explicit and teacher-directed instruction and targeted practice much of the time.

Sledge and Pazez (2013) examined special education teacher effectiveness within the context of the emphasis placed on teacher quality in the nationwide focus on teacher evaluation reforms. The authors reviewed reform efforts and practices implemented to distinguish between

effective and ineffective teachers and how those reform efforts have been applied to special education teacher effectiveness. Sledge and Pazez (2013) noted that efforts to measure teacher effectiveness have largely been focused on teacher evaluation systems that attempt to delineate teacher performance standards largely through observation protocols and other measures such as teacher portfolios, professional development goals, and measures of student progress. Sledge and Pazez (2013) contended, however, that current efforts aimed at assessing teacher effectiveness have largely been ambivalent to assessing special education teacher effectiveness. The authors were of the view that assessing features of effective special education teaching must take into the consideration the distinct roles and responsibilities required of special education teachers, variations that exist in teacher preparation programs for general and special education teachers, as well as the skills, knowledge, and dispositions that distinguish general and special educators. Sledge and Pazez (2013) were of the view that reform efforts to distinguish effective educators from ineffective ones will only be successful with special education teachers if the purpose, context, and roles and responsibilities of special education teachers are incorporated into the design and implementation of teacher evaluation measures.

Jones (2016) discussed what is expected of special education teachers as part of an examination into issues surrounding special education teacher evaluation. According to Jones (2016), the ideal expectation is that both general and special education teachers should be held to same standards in terms of teacher accountability reforms efforts. At the same time, the author noted that special education teaching is defined by roles and responsibilities that set effective special education instruction apart from general education. Jones (2016) argued that those unique roles and responsibilities that characterize special education teacher effectiveness should be factored into the current evaluation efforts and practices. To the author, special education

teaching is marked by three distinct features. The first is that special education instruction is planned and delivered in an individualized manner. This entails that special education teachers are able to deliver instruction aligned with students' individual needs, preferences, and abilities as mandated by their IEPs. In addition to delivery, special education teachers are expected to demonstrate expertise in organizing and coordinating supports and services that each student with disabilities needs to succeed. The second feature involves the unique instructional goals and professional responsibilities of special educators. According to Jones (2016), special education teaching should be focused on delivering targeted, direct, and explicit instruction and rooted in implementing evidence-based practices and interventions. The instructional responsibilities of special educators, which transcend instructional delivery, should also be considered in assessing their effectiveness. The author indicates the special education teacher spend a paltry amount of their time, 16% to be specific, on delivering instruction with the remaining quantum of time expended on a variety of responsibilities, including IEP development, curricula modifications, assessment development, and progress monitoring. The third and final feature has to do with the philosophical leanings of special education pedagogy. Jones (2016) observed that special education teaching is rooted in behavioral theory with its emphasis on prescriptive teaching, identification of processing deficits, and implementation of interventions to mitigate the effects of such deficits. This pedagogical approach serves as the basis for explicit instruction, which is seen as a fundamental feature of effective special education teaching. Jones (2016) further argued that the defining features of special education teaching, which set special education apart from general education, are at odds with the evaluation measures adopted to evaluate special education teacher effectiveness.

## **Current Evaluation Methods and Associated Challenges**

Current methods used to evaluate special education teachers and their associated challenges is perhaps the most prominent theme in the limited special education literature regarding the issue of special education teacher evaluation. According to Jones (2016), there is “long commentary surrounding the challenges of evaluating special educators” (p. 72). The number of studies under this theme is testament to this observation. The studies reviewed for this theme include Brownell and Jones (2015), Crowe, Rivers, and Bertoli (2017), Guartico (2016), Holdheide (2013), Holdheide, Goe, Croft, and Reschly (2010), Sledge and Pazez (2013), Benedict, Thomas, Kimerling, and Leko (2013), Johnson and Semmelroth (2014), Joyce, Harrison, and Murphy (2016), and Steinbrecher, Selig, Cosbey, and Thorstensen (2014). Almost all the studies commented on the methods that are currently being used to evaluate special education teachers and the challenges that emanate with those methods. Three studies, Crowe et al. (2017), Guartico (2016), and Holheide et al. (2010) were empirical, using mixed (Crowe et al., 2017), qualitative (Guartico, 2016), and quantitative (Holheide et al., 2010) approaches. The remaining studies (Benedict et al., 2013; Brownell & Jones, 2015; Holheide, 2013; Johnson & Semmelroth, 2014; Joyce et al., 2016; and Steinbrecher et al., 2014) were descriptive or conceptual.

Brownell and Jones (2015) examined research-based evaluation methods that have been adopted by states and school districts as part of policy reforms surrounding teacher evaluation and compensation reforms. The authors explored the methods to ascertain their validity and implications for use with special education teachers. Brownell and Jones (2015) framed their exploration by first discussing the purposes for evaluating special education teachers. According to the authors, special education teacher evaluation serves three purposes: (a) providing

information to states regarding the distribution of effective special education teachers across districts and schools, (b) helping school districts to plan targeted professional development efforts in terms special education teacher knowledge, instructional needs, and skills, and (c) identifying effective special education teachers for recognition while sorting out incompetent ones.

***Student Growth Measures.***

Brownell and Jones (2015) discussed current evaluation methods, including the research that supports their use, and identified issues that emanate from applying those methods to special education teachers. The researchers put the evaluation methods into two categories: measures of student growth and classroom observation protocols. The student growth measures identified by the authors were value-added models (VAM) and student learning objectives (SLOs). The authors noted that VAMs were the bedrock of evaluation reforms efforts, largely due to federal initiatives and reform efforts. VAMs are, however, problematic when used for special educators because of issues relating the number of students included in VAM calculations, the low scores of students with disabilities, and the challenge of attributing student growth to one teacher when special education instruction is nested across general and special education. For students with severe disabilities who are not able participate in general standardized assessments, SLOs have been promoted to assess the effectiveness of their teachers. Brownell and Jones (2015) acknowledged that SLOs offer several benefits over VAM in terms of being readily interpretable to teachers and administrators than VAM. In addition, SLOs provide a better measure of teacher impact on student learning. On the other hand, the researchers pointed out that there is little empirical evidence supporting the use of SLOs to reliably measure teacher effectiveness.

### *Classroom Observation Measures*

With regard to classroom observation measures, Brownell and Jones (2015) noted that most states have revamped their evaluation systems and adopted rigorous evaluation tools such as the FFT, Marzano's Teacher Evaluation (MET) tool, and the Classroom Assessment Scoring System (CLASS). Other measures, as discussed by the authors included stakeholder surveys, individualized education programs, and collaborative and effective teamwork tools. The authors admittedly highlighted the benefits of using the FFT in particular to evaluate special education teachers. The benefits included having a uniform vision of all professionals as a result of using a common evaluation tool, the FFTs alignment with teacher professional standards, as well as the cost-effectiveness resulting from employing a common observation tool for all teachers. On the contrary, the researchers noted that the FFT presents some challenges when used to evaluate special education teachers. One significant challenge was that districts do not typically use the domains 1 and 4, which comprise the planning and preparation and professional responsibilities domains. The authors contended that without these two domains, critical pieces of information that characterize effective special education practice are not accounted for. The FFT's constructivist orientation is also at odds with the social learning theory, behavior theory, or a combination of cognitive, social learning, and behavior theories upon which special education instruction is based. The authors concluded that there is limited research to demonstrate that the FFT, MET, and CLASS evaluation tools are able to distinguish between effective and ineffective special education teachers. Their validity for use in special education has not been researched and established. By the same token, Jones and Brownell further argued that the degree to which other measures such as stakeholder survey, individualized education programs, and collaborative and effective teamwork are effective in evaluating special educators has not been established,

although some states and school districts have modified their evaluation systems to include such measures.

Crowe et al. (2017) examined the absence of observation tools designed for use in special education classrooms, especially in self-contained settings. The researchers noted that special educators, regardless of their diverse roles and responsibilities, are being evaluated with observation protocols designed for general education settings, under the guise of maintaining equity in evaluation standards. Crowe et al. (2017), therefore, reviewed the components of the classroom context that evaluation tools designed for general education classrooms assess to identify components that were specific to special education contexts. The researchers analyzed a sample of nine observational tools using content and factor analysis. The researchers observed that the majority of the tools identified were designed for general education with only two geared towards special education. Crowe et al. (2017) found that most existing observation tools revolved around three constructs of teacher behaviors which were organization, instruction, and classroom climate but ignored the role of paraprofessionals as well as teacher professional behaviors. The researchers called for the design of specific evaluation tools for special education classroom that can account for shortfalls of existing systems and which can account for special educator professional behaviors as well as classroom interactions with support staff.

Sledge and Pazey (2013) considered the difficulties related to implementing emerging evaluation efforts in special education as part of a broader review of whether nationwide teacher assessment efforts recognize and account for the unique roles and responsibilities of special education teachers. Like other researchers, Sledge and Pazey (2013) argued that teacher evaluation efforts must be cognizant of inter and intra variations that exist between general and special educators. The researchers subsequently compared and contrasted the roles and



responsibilities of special educators to their general education counterparts, highlighting the merits and demerits of existing approaches used to measure teacher effectiveness and articulating the difficulties of applying those approaches in special education teacher evaluation.

Sledge and Pazey (2013) argued that any effective metric used to assess teacher effectiveness must be embedded in a clear understanding of professional roles and responsibilities teachers are expected to perform. The researchers noted that although the responsibilities that special education teachers are asked to assume are similar to general education teachers in some respects, they are different at the same time. Distinct responsibilities cited by Sledge and Pazey (2013) included special education teachers being asked to: collaborate between general education teachers and service providers, engage in frequent and sustained parental communication beyond what is expected of general educators, plan, design, and implement IEPs, have in-depth knowledge of special education mandates and policies, and supervise paraprofessionals.

Another distinct feature lies in how special and general educators are trained through teacher education programs. While all pre-service teachers are expected to graduate from teacher preparation programs with expertise required to ensure student success, the authors argued that special pre-service special education teachers are “expected to possess expertise in the distinct characteristics of various disability categories as well as the ways in which a particular student’s disability may manifest in different situations” (Sledge & Pazey, 2013, p. 236). In addition to the distinct responsibilities, Sledge and Pazey noted that special education teachers may assume roles and responsibilities that are contextually specific, such as different co-teaching models and variety of content mastery assignments, in order to meet student needs.

### **Special Education Teacher Responsibilities and Roles.**

Further confounding the above-mentioned difficulties, according to Sledge and Pazey (2013), are issues related to evaluator knowledge special education classrooms and the nature of expertise a special education teacher should possess, as well as issues surrounding the measurement of student performance. The researchers contended that some evaluators lack the requisite knowledge base regarding the special education classroom practices and procedures to able to perform accurate assessment of special education teacher performance. In terms of measurement of student performance as a component of special education teacher evaluation, Sledge and Pazey (2013) also criticized the use of VAMs as the measuring criteria. Their criticisms, specifically in connection with special education teacher evaluation, had to do with VAMs' inadequate data collection and calculations, incomplete data sets, small sample sizes, inaccurate calculations of value-added scores, varying testing conditions for students with disabilities, difficulties in assigning teachers to student test scores, and alternative assessments for students with severe disabilities. The researchers recommended modifications or replacement of existing evaluation instruments and the training of assessors to improve meaningful outcomes for special education teachers in the evaluation process. Sledge and Pazey also recommended that policy makers insist on involving teachers and administrators in the design and implementation of teacher evaluation systems for special education teachers to ensure that the complexities of special educator responsibilities and roles are accounted for in all components of the evaluation process.

Holdheide, Goe, Croft, and Reschly (2010) authored a research and policy brief under the auspices of the National Comprehensive Center for Teacher Quality that identified challenges in evaluating special education and English language learner (ELL) teachers with prevailing district

and state evaluation policies and practices. Through a nationwide survey and subsequent interviews, Holdheide et al. (2010) sampled over 1,000 respondents consisting of state special education directors from 36 states and 1,107 district level special education administrators for their study. Findings revealed disparities in approaches to teacher evaluation for special education and ELL teachers. One such disparity existed in the origin of the evaluation instrument itself. The researchers reported that the majority of instruments originated from school districts as 54% reported designing their own. It was significant that an overwhelming majority of school districts (68.8%) did not modify or adapt evaluation instruments or processes for special education teachers due to contractual restrictions, even though nearly half (49.9%) of respondents did not favor evaluating special education teachers with the same process as their general education counterparts. The researchers also noted disparities in terms of evaluators who had received explicit training for their roles. Although 62% of respondents indicated that training for evaluators was mandated, only 12.4% cited the existence of explicit training for evaluators.

To address the challenges identified in their findings, Holdheide et al. (2010) recommended the following: (a) include special education and ELL teacher in the design of evaluation systems, (b) identify common standards that characterize effective teaching for all teachers, including differentiated criteria where necessary, (c) integrate evidence-based practices for students with disabilities when designing evaluation models, (d) establish a culture of collaboration that builds trusts and empowers teachers, (e) incorporate concrete evidence of teachers' contribution to student learning, (f) ensure evaluation framework identifies and provides professional development needs of special educators, and (g) consider revisions to existing statutes or policies that restrict modifications to evaluation instruments for special educators.

Holdheide's (2015) study was a follow-up to the above findings and recommendations. The study noted that the existence of concerns relating to fairness, equity, validity and reliability, interrater agreement, measures of student growth, and legal ramifications had shifted the focus away developing evaluation systems that would address the unique responsibilities and roles of special education teachers. Holdheide (2015) was also concerned that the debate over whether or not separate evaluation systems should be adopted for special educators as opposed to a single system for all educators had the potential to derail the conversation and divert attention away from the years of work towards building a collaborative culture and shared responsibility for the education of students with disabilities. The concern, notwithstanding, Holdheide (2015) noted that quest to have evaluation systems that recognize unique roles and responsibilities is not peculiar to special education but to other groups of specialized educators as well. The study pointed to a few states and districts that had adopted strategies to address the unique roles and responsibilities of special educators and other specialized groups of educators. The common practices that emerged out of the strategies will be discussed under the next theme.

**VAMS.** The studies of Benedict, Thomas, Kimerling, and Leko (2013) and Steinbrecher, Selig, Cosbey, and Thorstensen (2014) both focused on measures of student growth in evaluations, specifically with VAMs. Both studies discussed why VAMs and associated practices are not appropriate for special education teachers. Benedict et al. (2013), however, differed from Steinbrecher et al. (2014), because it discussed VAM as part of a broader review of what special education teachers need to know about common evaluation methods used to evaluate their effectiveness, with the view of equipping special educators with the necessary skills and knowledge for the process. On the other hand, Steinbrecher et al. (2014) were focused solely on VAMs and argued against using VAMs to evaluate special educators due to their

potential shortcomings. The researchers were of the view that VAMs are problematic for special educator assessment due to state-level assessment practices for students with disabilities, mismatch between assessments and the abilities of students with special needs, and the diverse roles and responsibilities of special educators. In terms of state-level assessment practices, the researchers noted that the variation content when students with disabilities matriculate from one grade to another, misalignment of grade level assessments with actual student abilities, and the issue of accommodations during assessment administration were potential limitations that contribute to the disparity in performance of students with disabilities and their non-disabled peers. Further complicating the problem with VAM, according to Steinbrecher et al. (2014), are issues emanating from special educator roles and responsibilities, a concern shared by other studies (e.g., Benedict et al., 2013; Brownell & Jones, 2015; Holheide et al, 2010) reviewed in this chapter. Steinbrecher et al. (2014), however, identified three key considerations that affect VAMs scores for special educators. They included (a) the shared variance between general and special educators on one hand, and related service providers on the other, (b) class sizes in special education classes, which are typically smaller, raising concerns about shrinkage in calculation, and (c) non-random assignment of students with special needs to special educators, a situation that confounds student level factors that impact student growth. The researchers recommended the use of multiple measures such as IEP goals, portfolios, and student learning objectives (SLOs) to assess special educator effectiveness.

**SLOs.** Although SLOs have been promoted as one of the possible solutions to the difficulties associated to evaluating special education teachers, the study by Joyce, Harrison, and Murphy (2016) revealed that is it not a foregone solution. Joyce et al. (2016) examined the use of SLOs as a component of high-stakes evaluation within the context of special education

evaluation. As part of their studies, the researchers utilized data from a study on 19 states that participated in Race to the Top competitions. The researchers also gathered data from policy documents, implementation plans, technical documents, and interviews with state level education officials. The study found substantial variation in how SLOs were incorporated in evaluation scores. Joyce et al. (2016) specifically noted differences and lack of clarity among 18 states in terms of how SLOs were used for students with disabilities. The differences were especially evident in three pertinent decision areas: (1) the target population for SLOs, (2) the goals to be targeted in the SLOs, and (3) the assigned weight in determining final evaluation ratings. While acknowledging that the use of SLOs can have potential benefits for special education teacher assessment, Joyce et al. (2016) recommended that clear policy guidelines are needed to ensure clarity in implementation.

**Self-Contained Settings.** Guartico's (2016) study is one of the few empirical studies to examine the issue of special education teacher evaluation. The study solicited the perception of self-contained special education teachers with the view to understanding how they perceived current evaluation methods in terms of design, implementation, and relevance to the demands of their jobs. Guartico (2016) interviewed 28 teachers of students with severe disabilities placed in self-contained settings in the District of Columbia. The study found that special education teachers in self-contained settings were of the view that some components of existing evaluation systems were not applicable to special education. Special education teachers expressed the view that their responsibilities and roles such IEP-related roles were not recognized in evaluation rubrics. In contrast to the position of Brownell and Jones (2014) and Holdheide (2015) that a demand for a separate evaluation system for special education teachers will muddy the debate about improving special educator evaluation, Guaritco's (2016) study found that having a

separate system is viewed by special education teachers as necessary. Participants in the study also reported that evaluation scores are not necessarily dependent on implementing evidence-based practices of effective teaching in special education, but on authenticity, a special education teacher's relationship with evaluators, and evaluator skills in assessing special education teacher effectiveness.

### **Proposed Evaluation Approaches**

As the debate on how best to evaluate special education teachers rages on, researchers have proposed, among other things, designing measures that define effective teaching in special education (Holdheide et al., 2010), including clearly-stated expectations and performance criteria for special education classroom contexts (Sledge & Pazey, 2013). This section of the review examines efforts by researchers to design measures that are responsive to evidence-based practices and special education teacher evaluation needs.

Seven studies were selected for this theme. Five of the them (i.e., Barnes, Cipriano, Flynn, Rivers, & Xu, 2018; Doabler, Nelson, Kosty, Fien, Baker, Smolkowski, & Clarke, 2014; Elliot, Roach, & Kurz, 2014; Johnson, Crawford, Moyland, & Ford, 2016; Johnson & Semmelroth, 2014) discussed specific approaches to special education teacher evaluation. One study (Holdheide, 2015) discussed how some states and school districts had incorporated evidence-based practices in existing evaluation systems and the common themes that emerged from those systems. The final study (Israel, Kamman, McCray, & Sindelar, 2014) explored how evaluation can be embedded in the onboarding process by way of mentoring for beginning special education teachers. The studies that presented specific approaches differed in focus in terms of components of the evaluation process as well as category of special educators. Johnson et al. (2016) and Johnson and Semmelroth (2014) focused on the design of an observation tool

that can be used to assess the instructional special education teacher practice. Elliot et al. (2014) dealt with opportunities given to students with disabilities to access the general education curriculum while Doabler et al. (2014) examined the quality of interactions between special education teachers and students with disabilities. Barnes et al. (2018) presented the findings of an observation tool designed to evaluate instructional and social processes in self-contained classrooms.

### ***Model States and Districts***

Holheide (2015) reviewed practices and strategies adopted by four states and one large urban school district in designing or modifying existing evaluation systems to address the peculiar roles and responsibilities of special education teachers. The states and district included in the review were Colorado, Massachusetts, Pennsylvania, Ohio, and the Bartholomew County School Corporation in Columbus, Indiana. The selection of the states and school district, according to Holdheide (2015) were based on their inclusion of key instructional practices and strategies that recognized the unique roles of special educators and addressed student learning needs in evaluation models used. After analyzing the models, Holheide (2015) found five themes which included: (1) recognition that all teachers support the learning of students with disabilities, (2) promotion and reinforcement of evidence-based practices, (3) considerations given to measurement of student growth, (4) attention to specific roles and responsibilities, and (5) a focus on the vital role of administrators. Holheide (2015) concluded with the expectation that states and districts across the nation will learn from the ‘exemplary’ states and district discussed in the study by taking steps to align their evaluation rubrics with evidence-based practices and professional standards in special education.



### *Emerging Tools and Measures*

**RESET.** The studies of Johnson and Semmelroth (2014) and Johnson et al. (2016) both focused on the Recognizing Effective Special Education Teacher (RESET) observation tool. While the Johnson et al. (2016) study described the tool and the theoretical framework underpinning its critical components, the Johnson and Semmelroth (2014) study explained the actual design process. The RESET tool was the outcome of a partnership between researchers from the Boise State University in Idaho and the Idaho State Department of Education to develop an evaluation tool that ensured that special education teachers were integrated into the state's teacher evaluation and accountability system. The tool, according to Johnson and Semmelroth (2014) is situated within Domain 3, the domain used to assess instructional practices in Danielson's FFT rubric. The rationale was that the criteria used to assess special educators on the FFT did not provide specific, detailed, and actionable feedback that aligned with special education teachers' professional needs and evidence-based practices, although the authors determined that general components of the FFT were somewhat applicable to special education instructional practices. Johnson and Semmelroth (2014), therefore, defined the characteristics of an effective special education teacher upon which they premised their design. An effective special education teacher, according to Johnson and Semmelroth (2014), is one who is "able to identify a student's needs, implement evidence-based instructional practices and interventions, and demonstrate student growth" (p. 76). By using the RESET tool, Johnson and Semmelroth believed that special education teachers would be evaluated based on their use of evidence-based practices to improve outcomes for students with disabilities. It is worthy to note that the validity of the RESET tool yet to be fully determined and the researchers outlined the steps planned for the validation process. The validation process outlined by the researchers included steps to: (a)

determine the tool's reliability across times, raters, and teachers, (b) examine the results of RESET in comparison with similar measures of teaching efficacy, (c) determine the extent to which instructional ratings and measures of student growth correlate, and (d) examine the impact of feedback generated from RESET on instructional practice over time.

**RELATE.** Barnes et al. (2018) explored the use of the RELATE tool to examine special education teachers' use of instructional practices and interactions in self-contained classrooms with the view to collecting evidence of the tool's reliability and validity. According to the researchers, the RELATE tool was designed to fill a gap in special education teacher evaluation that remained unfilled even after the creation and validation of Johnson and Semmelroth's (2014) RESET tool. Barnes and colleagues argued that the RESET tool ignored social processes that are vital components of teaching and learning in special education classrooms. According to Barnes et al. (2018), the presence of teachers and paraeducators is a prominent feature in special education settings. Teachers and paraeducators collaborate to provide additional student services, manage lots of transitions, and deliver specialized instruction. As a result, the researchers believe special education teachers in such environments should not be evaluated with tools developed for general education settings. The researchers noted that using observation protocols adopted for general education settings in self-contained special education classrooms will obscure what actually transpires in those settings. Barnes et al. (2018) further noted that observers may not be able to accurately interpret and understand interactions between teachers and paraprofessionals with the proper evaluation tool.

Barnes et al. (2018) asserted that the RELATE tool "is the first empirically based tool that assesses quality of instruction and social processes in self-contained, special education classrooms serving primarily students with emotional and behavioral challenges" (p. 2). The

researchers adopted a mixed method approach, using content analysis of existing classroom observation measures, observations of special education classrooms, and interviews with a focus group comprising special educators, paraprofessionals, and administrators to examine the psychometric properties of the RELATE tool in self-contained settings. The study used videos collected from 47 classrooms, involving teachers of those classrooms and an average of two paraprofessionals per classroom from six school districts in the two northeastern states. To analyze their data, the researchers conducted generalizability (G) and decision (D) analysis. Findings from the study suggested that the RELATE tool has the potential psychometric and theoretical properties to be used to assess special education teachers and their self-contained settings. The study also found that there is a need for significant changes to at least one component of the tool.

**COSTI-M.** The study by Doabler, Nelson, Kosty, Fien, Baker, Smolkowski, and Clark (2014) was one of the few quantitative empirical studies on the issue of special education teacher evaluation in the special education literature. The study examined the effect of the *Early Learning in Mathematics (ELM)* curriculum on teachers' use of explicit instruction in the core academic area of math. The researchers wanted to measure the quality and intensity of explicit instruction as well as the implementation fidelity of the ELM curriculum through the use of a multifaceted observation system comprised of four instruments. The first instrument, the Classroom Observations of Student-Teacher Interactions of -Mathematics (COSTI-M) was used to capture the quantity of explicit instructional interactions between teachers and students. To measure the quality of explicit instruction, the researchers designed two other instruments namely: the Quality of Classroom Instruction (QCI) and the Ratings of Classroom Management and Instructional Support (RCMIS). The fourth and final instrument, the ELM Fidelity of

implementation instrument was used to assess the fidelity with which teachers implemented the ELM curriculum. Using these instruments, Doabler et al. (2014) conducted a total of 379 observations in 129 classrooms with about 2700 students across 46 schools in Oregon and Texas. Sixty-eight (68) of the classrooms were randomly assigned to a treatment group that utilized the ELM curriculum and teachers evaluated on the four instruments while 61 classrooms were assigned to the comparison group that utilized standard district curriculum. The study found that special education teachers who used the ELM curriculum delivered significantly higher rates of individual and group responses compared with the control group. The findings from the study suggest that an observation system based on evidence-based practices in special education might be useful.

**MyiLOGS.** Elliot, Roach, and Kurz (2014) described another approach known as the *My Instructional Learning Opportunities Guidance System (MyiLOGS)* evaluation system for special education teachers. The system is based on the opportunities to learn (OTL) framework that measures how teachers apportion instructional time, the content covered, that the quality of instructional practices, to create learning opportunities teachers for students. According to Elliot et al. (2014), the system was developed using findings from professional development research. To use MyiLOGS, teachers are required to take a qualifying test to document an individual teacher's knowledge of curriculum, instruction, and assessment practices that contribute to student learning. According to the researchers, the test is administered online and composed of 35 questions with multiple-choice, true or false, and short answer options. Teachers must attain a score of 85% to be considered proficient. Teachers then self-report daily instructional time, curricula content covered, and instructional practice within a lesson into the MYiLOGS system. A third component of the system called My Instructional Observation System (MYiOBS) is used

to collect observational data to ascertain the reliability of teacher's self-reported data. Teachers receive detailed feedback generated on their instructional practices every 40 days. Finally, an instructional plan growth plan that documents a teacher's progress on one or more instructional practices within the school is created.

According to Elliot et al. (2014), all the elements in MyiLOG system have been used in both general and special education settings, but are yet to be used in a formal evaluation system. In addition, the system has some notable limitations including issues relating to time needed for implementation, reliability of self-reports from teachers, framework needed to summarize and evaluate the culminating evidence to determine teaching effectiveness, and how to build student achievement component into the portfolio.

**Mentoring.** The study by Israel, Kamman, McCray, and Sindelar (2014) focused on special education teacher evaluation as a component of new teacher induction for beginning special education teachers. According to Israel et al. (2014), the research base on how to combine mentoring and evaluation of beginning special education teachers is sparse. The researchers noted that there were mixed findings within the limited literature with some researchers reporting negative effects of merging evaluation practices into mentoring programs while other studies pointed to the usefulness of teacher evaluation that provides formative feedback for new special education teacher development. Israel et al. (2014) therefore sought to investigate how one urban school district in Midwestern United States merged evaluation into its mentoring program to address the professional and emotional needs of new special education teachers. Specifically, the researchers were interested in the answers to the following research questions: (1) Within the context of a mentoring program with a strong evaluation component, what types of professional and emotional supports are provided for new special education

teachers? and (2) What is the relationship between professional and emotional supports within this mentoring program? Participants included 16 new special education teachers and five special education mentors. Data sources for the study were evaluation records, mentor records, and interviews with new special education teachers. Israel et al. (2014) found that evaluation guided feedback mentors gave to the new teachers as mentors provided detailed and concrete feedback to help the new teachers improve their practices relative to evaluation indicators. The study further found that emotional and professional supports were interconnected as mentors provided emotional supports are part of professional supports. Another significant finding was that new teachers had favorable views of evaluation as a component of mentoring experience. Although this study did not include the design of a specific evaluation instrument or tool as was in the case of the studies reviewed above, it provides a useful framework to guide emerging approaches to evaluation special education teachers and the role of peer mentors in the process.

### **Status of Research on Stakeholder Perceptions**

Teachers, administrators, and teacher educators have been identified as important stakeholders in the evaluation process (CEC, 2012; Holdheide et al., 2010; Woolf, 2015; Woolf, 2018). This section of the review sheds light on the current status of research into the perceptions and perspectives of the different stakeholders involved in developing and implementing, or impacted by evaluation reforms.

Seven studies (Glowacki & Hackman, 2016; Guartico, 2016; Lawson, 2015; Lawson & Knollman, 2017; Steinbrecher, Fix, Mahal, Serna, McKeown, 2015; Woolf, 2015, Woolf, 2018) were selected for review under this theme. Three studies (Glowacki & Hackman, 2016; Lawson & Knollman, 2017; Steinbrecher et al., 2015) focused on administrator perceptions, one (Lawson, 2015) examined administrator and special education teacher perspectives, and two

(Woolf, 2015; Woolf, 2018) explored the viewpoints of administrators, teacher educators, and special education teachers. Just one study (Guartico, 2016) was solely dedicated to exploring special education teacher perceptions. One notable finding is that with the exception of Woolf (2015) and Woolf (2018), all the studies involved participants (administrators and teachers) from the elementary level.

### ***Administrators Perceptions***

As indicated above, three studies (Glowacki & Hackman, 2016; Lawson & Knollman, 2017; Steinbrecher et al., 2015) examined administrator perceptions of the evaluation process although they differed in their respective areas of focus. Glowacki and Hackman (2016) explored the extent to which elementary administrators perceived existing evaluation systems as effective in addressing special education teacher roles and responsibilities, as well as perceptions relating to administrator proficiency in evaluating special educator performance. Lawson and Knollman (2017) also studied administrator beliefs regarding ability to evaluate and provide meaningful feedback but did not solicit views about special educator responsibilities and roles in relation to evaluation. Steinbrecher et al. (2015) on the other hand explored the views of elementary principals in terms of how their perceptions of what is important in special education matched with standards as set by the CEC.

Glowacki and Hackman (2016) surveyed 330 elementary level principals to investigate their perceptions relating to the effectiveness of the evaluation process as well as how the principals perceived their own competence in assessing special education teachers. Through basic and inferential statistics, Glowacki and Hackman conducted quantitative analysis of the survey data. Findings revealed that principals perceived existing evaluation systems in their districts as effective in evaluating special education teacher roles and responsibilities. In terms of

principals' perceived competence in evaluating special education teachers, however, significant differences were found between principals who had special education backgrounds and those who did not. Principals with special education knowledge reported being more effective in providing relevant feedback to special education teachers. On the other hand, principals who were not certified in special education reported feeling less effective in providing specific feedback on special education professional responsibilities and roles.

Lawson and Knollman (2017) also investigated administrators' perspectives concerning their ability to evaluate special education teachers. Unlike Glowacki and Hackman (2016), however, Lawson and Kollman's study was restricted to a sample of elementary principals who reported having no background in special education. Three administrators from a school district in southern California were interviewed regarding their: (a) experience in evaluating special education teachers, (b) perspectives in terms of the quality and quantity of training provided by their school district and through rating sessions, and (c) beliefs in their capacity to provide valuable feedback to special education teachers on their instructional practices. The interviews lasted about 45 minutes long and were analyzed using qualitative methods. Four major findings were reported for the study. First, administrators reported that they had not received any special education training before or during their time as evaluators in their school districts. The administrators reported that their experiences as evaluators, both in general and special education settings, compensated for their lack of special education background knowledge. Second, the administrators expressed confidence in their ability to evaluate and provide feedback despite their reported lack of background knowledge. This finding was in contrast with that of Glowacki and Hackman who found that administrators with no prior special education knowledge felt inadequate to evaluate special educators. Third, the administrators interviewed believed that the



parameters of good instruction in both general and special education should be held to same set of teaching and practice standards with no distinction between general and special education. In effect, they believed that “good instruction is good instruction regardless of subject matter or instructional setting” (Lawson & Knollman, 2017, p.33). Finally, the researchers reported that participants felt evaluation measures in their districts were designed for universal use and did not address the specific needs or contexts of sub-groups like those of special education teachers. The administrators admitted finding it more challenging to evaluate special education teachers in self-contained settings where students with severe disabilities were served.

Steinbrecher et al. (2015) also investigated administrator perceptions but their study differed from the above studies in its purpose. The researchers in this study wanted to determine the skills and knowledge administrators identified as important for special education teachers to possess relative to their instructional practices. The researchers also investigated how the identified skills aligned CEC initial preparation standards. Participants were five elementary level administrators from an urban district in southwestern United States who were interviewed for 20 to 45 minutes. Interview data were analyzed through constant-comparative qualitative methods. Findings from the study indicated that administrators expected special education teachers to be knowledgeable in special education policies and procedures such as meeting IEP goals and correct implementation of special education mandates like the Individuals with Disabilities Education Act (IDEA). The administrators also expected special education teachers to possess skills in the areas of collaboration, classroom management, differentiating learning for individual students, content area knowledge, instructional planning and strategies, and assessment. In terms of the alignment of the expectation to CEC standards, Steinbrecher et al. reported that administrators lacked knowledge on how to operationalize evidence-based practices

in a special education classroom and were more concerned about delivery of curricula content than on meeting the individual student needs. The researchers also found that administrators emphasized dispositions more than specific instructional strategies for effective teaching in special education settings, an indication of their lack of knowledge of evidence-based special education practices to inform reliable evaluations.

### ***Multiple Stakeholders***

The two studies of Woolf (2015) and Woolf (2018) had similar design to the Steinbrecher et al. (2015) study. Woolf's (2018) study was an extension of the earlier (Woolf, 2015) one. Both studies explored whether stakeholders viewed CEC professional special education standards as important in measuring special education teacher effectiveness. The studies differed from Steinbrecher et al.'s (2015) study in terms of the sample of participants who were drawn from a triad of administrators, special education teachers, and teacher educators for both studies. In the first study, Wolf (2015) surveyed 238 participants comprising 127 special education teachers, 58 administrators, and 53 teacher educators drawn from Northeastern United States to ascertain whether participants perceived the professional skills subsumed within CEC's national standards to be important for special education teacher effectiveness. Wolf (2015) also investigated which skills appeared more or less important for special education teacher effectiveness, and whether ratings of importance differed relative to stakeholders' distinct professional roles. Data were analyzed using inferential statistics such as ANOVA, MANOVA, and *t* tests. Results indicated that stakeholders agreed on the importance of the CEC standards. The study also found that stakeholders' ratings regarding the which skills were more or less important were similar across groups. In terms of relative importance of skills, however, Wolf (2015) noted differences. School administrators rated four out of five domains higher than special educators and teacher educators.

Woolf's (2018) study built on the findings of the 2015 study. Wolf investigated which skill domains in the CEC standards were identified as vital for special education teacher effectiveness across and within the stakeholder groups. Data collected for the study were subset data from the earlier (Woolf, 2015) survey administered to participants. As part of the survey, participants were asked to respond to open-ended prompts that asked them list skills they perceived to be critical to special education teachers' effectiveness, but which were not represented in the survey questions. Of the 238 participants in initial study, 140, consisting of 76 special education teachers, 33 school administrators, and 31 teacher educators provided responses that were analyzed for the study, using qualitative methods.

Results showed that stakeholders viewed three major skills as critical to special education teacher effectiveness. The skills were: (a) understanding disability and associated impact on learning, (b)integrated expertise, and (c) instructional flexibility. Understanding disability and associated impact on learning involved specialized knowledge about disability that is critical for special education teachers to fulfill their roles and responsibilities such as ability to identify individual learning needs, differentiate learning for students with disabilities, and build trusting relationships. Integrated experience required that special education teachers demonstrate proficiency and relevant content area knowledge, together with pedagogies and evidence evidence-based instructional practices. Regarding instructional flexibility, the study found that stakeholders expected special education to utilize multiple strategies when working with students. This skill demanded that special education demonstrate fluidity and immediacy in instructional delivery and accommodation provision or modification based on data. In terms of which skills emerged as critical for special education teacher effectiveness relative to participant groups, the study found stakeholders perceive some critical skills differently. For instance,

Woolf (2018) reported that special education teachers and school administrators emphasized professional preparedness, which had to special educators' ability to fulfill nonteaching roles, but teacher educator did not. School administrators and special educators also stressed advocacy, and role flexibility as critical skills. Teacher educators, however, did not emphasize those skills in the same way. On the contrary, teacher educators rated ongoing reflection and professional development as critical skills. The findings suggest that special education teachers, administrators, and teacher educators have differing expectations with regard to critical skills effective special education teachers must possess.

### ***Special Education Teacher Perceptions***

Guartico (2016) and Lawson (2015) were the only studies that explored special education teacher perceptions of the evaluation process. Lawson's (2015) study, however, examined administrator perceptions as well. Both studies used qualitative approaches although Lawson (2015) utilized quantitative methods as well to examine administrator beliefs.

Lawson (2015) studied special education teacher perception of the RESET observation tool, which, as alluded to earlier, is designed for the purpose of observing instructional processes in special education settings. It must be pointed out that the RESET tool is yet to be adopted commercially for mass use across states and districts in comparison with tools such as the FFT, CLASS, or Marzano's Teacher Evaluation Model. For the study, Lawson (2015) interviewed five special education teachers from California with the goal of exploring the following four areas: (a) their experiences relating to observation and evaluation by an administrator, (b) their beliefs regarding the appropriateness and validity of the RESET rubric items in evaluating their instructional practices, (c) suggestions for improving instructional practice of special educators,

and (d) suggestions for improving rubric items on a special education observation tool. Lawson (2015) conducted a thematic analysis of the interview data resulting in four themes.

The first theme indicated that special education teachers believed administrators should do more by way of providing useful feedback to inform their professional growth instead of simply going through motions of observation and evaluating. Evaluation, according to the teachers, should be a formative process infused with ongoing support and mentoring. The second theme revealed that participants were divided in terms of their perceptions of fairness of the evaluation process. While some participants felt that the process was fair, others held contrary viewpoints. Those who perceived the process as unfair cited the non-teaching professional responsibilities such as the IEP-related responsibilities, communication with parents, and service coordination as tasks not reflected and counted for final ratings. Participants who perceived the process as fair viewed it as such only from the instructional component. The next theme showed that special education teachers viewed good instruction as easily discernible regardless of the evaluation instrument used though they also suggested that special education instruction sometimes appears different from general education. Ironically, participants could not articulate the distinguishing features of good special education teaching. The final theme found that participants desired more observations from administrators instead of the two that they received annually in their districts. Participants especially favored informal visits, which they perceived would afford administrators the opportunities to identify areas of improvement and growth. The findings from the study suggest that special education teachers were more concerned with the frequency of visits and the nature of feedback they received from administrators as opposed to the specifics of the evaluation rubrics.

The study by Guatico (2016) also examined special education teachers' perceptions of the design and implementation of evaluation systems, without reference to a specific observation tool as was in the case of Lawson (2015). The purpose of Guatico's (2016) study was threefold: (a) to explore the perception of special education teachers in self-contained elementary settings regarding the implementation of current evaluation models, and (b) to provide teachers the opportunity to share their views on evaluation practices relative to their professional responsibilities, and (c) to solicit recommendations for teacher evaluation relative to special education teacher preparation and professional development needs. Twenty-eight participants were selected from three school districts in the Washington DC metropolitan area. Two out of the three districts from which participants were selected used their own district-designed observation measures while the third district utilized the FFT. Data were collected through interviews and analyzed using quantitative methods. Guatico (2016) reported four themes relating to issues of: relevance, training, feedback received, validity.

On issues of relevance, the study found that participants were of the view that certain components of the evaluation rubrics did not apply to settings in which they teach or the evidence-based practices used. These sentiments were especially expressed by participants from the district that used the FFT for evaluation purposes. Participants from the other districts with locally-designed systems had positive views of their systems and their rubrics. Another issue of relevance was that participants showed indifference towards the evaluation process and viewed it merely as another item on their to-do list. On the second theme, issues of training, participants reported receiving training, but not specific to their teaching programs. The majority also perceived their administrators as lacking adequate training to evaluate their practice, but they did not view peer evaluators or expert teachers the same way. Issues relating to feedback received

showed that participants perceived feedback as minimal. Finally, participants were of the view that authenticity, relationship with assessors, and scoring issues impacted the validity of the evaluation process. Regarding suggestions to improve the evaluation process, Guartico (2016) found that participants favored changes to evaluators, preferring specialists and peer evaluators who are familiar with special education settings. Participants also recommended changes to how they are observed, encouraging the use of informal observation, parent surveys, using videotapes, and student surveys. Participants also suggested changes to the performance indicators, and evaluation and scoring component of the process. In addition, Guartico (2016) found that special education teachers strongly recommended the need for evaluators to have a background in special education teaching and contexts.

### **Discussion**

The purpose of the literature review was to unearth what has been written in the special education literature regarding the issue of special education teacher evaluation and to use the findings to guide the methods for the study. The findings indicate the need to clearly define what constitutes effective teaching in special to guide the development or modification of evaluation systems to address the needs of special education teachers (Jones & Brownell, 2014; Sledge & Pazez, 2013; Jones, 2016). The limited literature on what constitutes effective special education teaching reveals that special education teaching is distinguished by explicit, repeated, and direct instruction coupled with guided practice. This underlying feature is at odds with the constructivist philosophy underpinning most evaluation systems, designed for use with general education. Although there is ongoing research to incorporate the core features of effective special education teaching into existing systems such as the FFT (Brownell & Jones, 2015; Holheide, 2015) and new tools like RESET (Johnson & Semmelroth, 2014), little is known about efforts to

do so with other observation measures such CLASS or Marzano's framework. We also do not know how successful those efforts have been at addressing concerns about equity and fairness for special education teachers.

It is apparent from the review that there are a multiplicity of approaches to evaluating special education teachers both in terms of classroom observation and student performance measures. One obvious fact, however, is that no approaches have been validated for use with special educators (Brownwell & Jones, 2015, Johnson & Semmelroth, 2014, Jones & Brownell, 2014, Sledge & Pazey, 2013). Researchers have expressed the need to close the achievement gap between students with disabilities and the non-disabled counterparts and improve post-school outcomes. It is therefore imperative to ensure that teachers who serve students with disabilities are of high quality. The unique responsibilities and challenges associated with evaluating special education teachers require meticulous consideration and need for a review of existing teacher evaluation systems to accommodate their needs (Sledge & Pazey, 2013). For evaluation systems to address the needs of special education teachers, they must include multiple measures that can capture the multiple roles and responsibilities of the special education teachers, specific to special education classroom contexts, based on evidence-based practices, and address student outcomes and teacher professional development needs (Brownell & Jones, 2015, Holdheide, 2013; Holheide et al., 2010, Johnson & Semmelroth, 2014; Jones, 2016; Sledge & Pazey, 2013).

One consistent theme relating to the issue of measuring student growth was the use of VAMs for special education teachers. The inappropriateness of VAMs as measures of special education teacher effectiveness is thoroughly discussed in the special literature. With the 2015 passage of Every Student Succeeds Act (ESSA) that reduced federal influence in the evaluation and accountability requirements, however, the role of VAMs appears to be shifting. In an effort



to ascertain the extent to which states have revised their assessment and teacher evaluation plans to take advantage of the flexibility offered by ESSA, Close, Amrein-Beardsley, and Collins (2018) analyzed 51 ESSA plans submitted to the US Department of Education by states and interviewed state level education personnel for information regarding their current teacher evaluation systems. Close et al. (2018) found that although states have not fundamentally altered their student assessments, the role of growth models such as VAMs in teacher evaluation is gradually changing. According to the Close et al. (2018), “language about holding teachers accountable for their value-added effects, or lack thereof, is less evident in post-ESSA plans” (14). Granted, there are still some states that encourage use of VAMS but Close et al. (2018) noted that even in those states, VAMs are offered as “off-the-shelf options” (p. 13) for school districts that may not have the resources to readily design their own evaluation models. In place of VAMs, states are shifting to using multiple measures and formative, instead of summative evaluations. While the trend appears encouraging, we do not yet know the extent to which special education teacher needs were factored into post-ESSA evaluation plans.

Researchers have offered some recommendations on special education teacher evaluation. Some have called for modifications of existing systems and rubrics, such as the FFT, to better account for special educators (e.g. Brownell & Jones, 2015; CEC, 2012, Holdheide et al, 2010, Johnson & Semmelroth, 2014). Other researchers have advocated for designed measures specifically for special education teachers (Barnes et al., 2018; Crowe et al., 2017; Sledge & Pazey, 2013). Other recommendations include the provision of adequate training for assessors and principals to help them develop the required expertise needed to assess accurately and to provide useful feedback peer observations, evidence-based measures, and involvement of special educators in designing evaluation systems (Sledge & Pazey, 2013). While approaches

and tools like RESET, RELATE, MYiLOGS, COSTI-M have been developed by researchers to target special education teaching and the contexts in which special educators work, research is needed to assess the validity of these measures. In addition, most of the measures have been limited to the local areas where they were developed or to specific components and settings of special education teaching.

The review also reveals large gaps in the literature relative to special education teacher perceptions and attitudes towards evaluation systems. There is sparse empirical literature on special education teacher effectiveness and evaluation. Of the 28 studies selected for review, only 11, comprising five qualitative (i.e. Guartico, 2016; Israel et al., 2014; Lawson & Knollman, 2017; Steinbrecher et al., 2015; Woolf, 2018), four quantitative (i.e. Barnes et al., 2018; Doabler et al., 2014; Glowacki & Hackman, 2016; Woolf, 2018), and two mixed methods (i.e. Crowe et al., 2017; Lawson, 2015) were empirical. The remaining 17 studies (60%) were composed of policy briefs and conceptual or theoretical pieces (e.g. CEC, 2010; Holdheide, 2013; Johnson & Semmelroth, 2014; Jones & Brownell, 2014). The literature on special education teacher perceptions, attitudes, perspectives, and opinions is even sparser. Just two studies (Lawson, 2015; Guatico, 2016) explored special education teacher perceptions. Guatico's (2016) study used basic qualitative methods and sampled participants from elementary self-contained settings. Lawson's study partially examined teachers' perception as part of a broader study with administrators and special educators as participants. There is an obvious need for research involving multiple categories of special education teachers in diverse settings and school levels to fill the gaps on special education teacher perceptions and attitudes toward evaluation systems. Empirical studies on difficulties associated with special education teacher evaluation are also needed. There is also an urgent need for research to validate effective

instructional practices in special education as well as systems that can be used to evaluate special education teachers. (Brownwell & Jones, 2015, Johnson & Semmelroth, 2014, Jones & Brownell, 2014, Sledge & Pazez, 2013). In addition, the findings demonstrate the need to study special education teacher perceptions and recommendations that can be used to inform the development of evaluation systems. The present study is intended to explore how special education teachers perceive evaluation systems with the intent of filling some of the identified gaps and contributing to the limited research on the issue.

### **Summary of Chapter**

This chapter reviewed the literature of the issue of special education teacher evaluation. A total of 28 studies met the inclusionary criteria and were selected for the review. A comprehensive and systematic reading of the studies culminated in four major themes: (i) issues relating to defining what is considered effective teaching in special education, (ii) current evaluation methods for special education teachers and associated challenges, (iii) proposed evaluation approaches and measures, and (iv) the status of research on stakeholder perceptions of the evaluation process. The findings from the literature synthesis will be used to inform the next chapter.

## **CHAPTER THREE: METHOD**

### **Introduction**

The purpose of the study was to explore special education teachers' perceptions relative to the evaluation systems used to assess their performance and effectiveness. Specifically, the study explored perceptions pertaining to the appropriateness and ability of existing evaluation measures to distinguish the multiple responsibilities and roles, contexts, pedagogical and professional development aspirations of special education teachers. The review of relevant literature on the topic and subsequent findings informed the choice of methodology for this study. As evident in the preceding review, much of the limited research on the topic has largely coalesced around commentary on existing evaluation methods and resulting challenges relating to assessing special education teacher performance (e.g., Brownell & Jones, 2015; Crowe et al., 2017; Holdheide et al., 2015; Sledge & Pazey, 2013, Steinbrecher et al., 2014) as well as issues relating to the definition and determination of features of effective special education teaching (e.g., CEC, 2012; Johnson, 2015; Jones, 2016; Jones & Brownell, 2014). There is also limited effort to design measures such as Johnson and Selmmelroth's (2014) RESET tool and Barnes et al.'s (2018) RELATE measure; tools that are considered responsive to special education teachers' needs and the contexts within which they work. The literature review further revealed a conspicuous absence of research on teacher perceptions about the evaluation process and the measures used, in spite of the importance of teacher inputs and calls from researchers to include teacher voices in the design and implementation of the evaluation process (CEC, 2012; Guartico,

2016; Jiang, et al., 2015). The two studies that focused on special education teacher perceptions (i.e. Guartico, 2016; Lawson, 2015) were limited to special education teachers at the elementary school level and utilized qualitative methods with small sample sizes. The study fills some of the gaps and contribute to the limited, yet growing, research regarding special education teacher perceptions through a quantitative study that sampled participants who are reflective of the heterogeneity of special education teacher responsibilities and roles at the secondary level. The choice of secondary level teacher participants is significant in view of the fact that secondary school marks a critical milestone, where students with disabilities are expected to receive transition services, as mandated by IDEA, to facilitate better academic, social, and post-school outcomes.

The study used Fullan (2001) and Doyle and Ponder's (1977) teacher perception of change conceptual framework to underpin the exploration of special education teacher perceptions about teacher evaluation systems. The following research questions guided the study:

- How do special education teachers rate their understanding of the teacher evaluation system used in their district?
- What are the perceptions of special education teachers about teacher evaluation systems with regard to their multiple roles, responsibilities, and contexts?
- What are the perceptions of special education teachers about evaluations systems with regard to the ability of these systems to address professional development, practice, and pedagogical needs?
- Do special education teachers' perceptions of teacher evaluation systems vary by:
  - a) job categories
  - b) gender, and
  - c) years of experience.

## **Methods and Procedures**

### **Methods**

Quantitative methods were used to explore the research questions of this study. According to Check and Schutt (2012), quantitative studies collect numbers and quantities as basic data and utilize an array of statistical procedures to analyze the data. Examples of methods used to collect data in quantitative research include surveys and experiments. Quantitative methods are used when the intentions of the research are explanation, description, or evaluation (Check & Schutt, 2012). Data are collected on predetermined instruments that yield statistical data (Creswell, 2003).

Quantitative research may also be classified as either experimental or non-experimental (Ary, Jacobs, & Sorensen, 2010). In experimental research, variables are manipulated to study the effect of one variable on another. Subjects are randomly assigned to treatment or control groups. Random assignment gives each subject an equal and independent chance of being assigned to any group, devoid of the researcher's judgment or bias. On the other hand, non-experimental research does not involve random assignment of subjects or the manipulation of variables. Instead, the researcher "identifies variables and may look for relationships among them" (Ary et al., 2010). Examples of non-experimental include correlational research and survey research (also known as descriptive research). In survey research, for example, the researcher, through the use of questionnaires or interviews, collects numbered data and analyzes them using statistical methods, to describe trends and to test research questions or hypotheses. Meaning of data is interpreted by connecting the statistical findings back to past research on the issue under study (Creswell, 2015).

For the purpose of this study, which is focused on studying the perceptions of different categories of special education teachers across multiple settings and school levels, survey

research was used. According to Ary et al. (2010), survey research allows the researcher to summarize the characteristics of different groups or study their attitudes and perceptions towards an issue. Survey research is useful when the researcher is interested in describing trends, determining people's opinions regarding policy issues, and identifying important beliefs and attitudes of individuals. Survey research can also unearth helpful information for program evaluation in schools (Creswell, 2015).

Like any quantitative study, independent and dependent attributes were identified, measured, and studied. A variable is defined by Creswell (2009) as a "characteristic or attribute of an individual...that can be measured or observed and that varies among people...being studied (p. 50). Variables are typically classified into independent and dependent variables. An independent variable is an attribute or characteristic that is hypothesized to influence or vary the outcome of another variable. The dependent variable, on the other hand, is the attribute or characteristic seen to be dependent or under the influence of the independent variable (Check & Schutt, 2012; Creswell, 2015). The independent variables for this study were special education teacher job categories, gender, years of experience, and familiarity with existing teacher evaluation system. The dependent variable was the special education teachers' perceptions of the evaluation system used in their district.

### **Instrument**

As indicated above, the study employed a quantitative design based on descriptive and inferential statistics and was not experimental in nature. In designing instruments for survey studies, Creswell (2015) recommended that researchers first consider whether there are available instruments that can reliably measure the proposed variables. Researchers can also consider modifying an existing instrument before deciding to design their own instrument. For this study,

a Likert-type instrument was adopted and administered to all participants. The researcher adopted an instrument used by Jiang et al. (2015) in their study of teacher perceptions of the Chicago Public Schools' evaluation system. A request for permission to use the instrument was sent to the authors and approval was granted. With their approval, the researcher modified portions of the original instrument to align with the purpose and research questions of the study, and with the intent of getting a valid instrument that would be able solicit responses needed to answer the research questions. The modified instrument was then tested for validity and reliability through a pilot study the researcher conducted. The instrument was divided into three categories in harmony with study's conceptual framework: *clarity*, *practicality*, and *cost*. Items on *clarity* dealt with special educators' understanding of the evaluation rubric and processes. Items on *practicality* were further delineated into measures of *evaluator perceptions*, *feedback*, *professional development*, and *student growth*. The last category of items was classified as *cost*, and elicited responses pertaining to special educators' perceptions in terms of the stress, frustration, and efforts associated with the evaluation process.

**Validity and Reliability.** Individual-level reliabilities on the original scale ranged from .68 to .87 (Jiang et al., 2015). Following the adoption and modification of the instrument, the researcher solicited the assistance of expert reviewers made up of doctoral students and faculty to ensure validity and reliability. The researcher also conducted a pilot test of the instrument through a pilot study with a sample of 10 participants drawn from a population similar to the target population and sample frame. The pilot study revealed that at least two items needed revision. Revisions were made to items that participants had difficulty answering or skipped most and to items the expert reviewers raised questions about. Participants in the pilot study were not part of the final sample for this study since they provided feedback on the instrument



(Creswell, 2015). Reliabilities of final scales were also measured by Cronbach's alpha as part of the data analysis. The scores ranged for a low of .70 to high of .89, all within acceptable and good reliability range.

### **Population and Sampling**

Survey research techniques require that researchers clearly identify and define the population, sampling frame, and sample of the study (Ary et al., 2010, Check & Schutt, 2012; Creswell, 2015; Gorves, Fowler, Couper, Lepkowski, Singer, & Tourangeau, 2009). A population is defined as the group of individuals with one characteristic that differentiates them from other groups. This is the group of individuals to which the findings of the study may be generalized (Check & Schutt, 2012; Creswell, 2015). The sampling frame, also referred to as the population frame (Groves et al., 2009), is a set or unit of the target population from which a probable list of sample survey will be selected. Finally, the sample is a group of participants who are selected from the sample frame. This is the group that is studied and from which data and measurements is sought (Creswell, 2015; Groves et al., 2009).

For this study, all special education teachers in the United States constituted the population. All special education teachers in a southwest Florida school district comprised the sampling frame. From this frame, a sample of special education teachers were selected for the study. All special education teachers at the secondary level (i.e. middle and high schools) in the selected southwest Florida school district were included in the sample. The school district was selected because of its use of the Marzano Framework to evaluate its teachers, including special education teachers. As revealed in the literature review, there are a handful of studies that have examined the appropriateness of using evaluation tools such the FFT for special education teachers. However, no study has been done so with the Marzano Framework. For the purpose of

this study, a special education teacher was defined as a teacher who possesses a Florida Exceptional Student Education (ESE) K-12 certification and who has the primary responsibility of teaching students with students with disabilities. That included special education teachers in general education, resource room, self-contained settings, and special schools or centers. General education teachers with ESE-K12 certification or supporting staff and service providers such as behavior specialists, speech therapists, occupation therapist, and physical therapists were not included in the sample.

To determine the minimum sample size needed for a study, Ary et al. (2010) indicated that the following four elements are needed: (1) the statistical test to be used, (2) the acceptable probability of type I error (alpha level), (3) the effect size that separate significant from trivial, and (4) the desired probability (power). At times, however, it is possible to study the entire population, especially when it is small and members of the population can be identified without much effort. This type of survey, known as census study, allows the researcher to draw conclusions about the entire population. Creswell (2015) pointed out that random sampling, hypothesis testing, and use of inferential statistics for sample selection are not necessary for census surveys. The study used the census survey approach, and therefore, did not conduct inferential statistical calculations to arrive at a specific sample size.

The selection of participants of the study begun with an application to the school district for permission to recruit special education teachers at the secondary level for this study. The office responsible for accepting, reviewing, and approving research applications requires all individuals/researchers who wish to conduct research within the district to submit an electronic application and obtain approval before proceeding. The school district defined educational research to include any data collection from or about the district staff, parents, students, and

departments. In addition, the district required that all applications must be accompanied by Institutional Research Board (IRB) approval letter from the sponsoring institution, which the case of this was the University of South Florida IRB. With the school district’s permission and final approval for the university IRB, all special education teachers from all the 21 middle, 18 high, and 5 exceptional schools in the district were surveyed for the study.

**School District Profile.** The selected school district is one of the top 10 largest districts in the southeast with over 100,000 students. It is also the largest employer in the county where it located, employing more than 16,000 full-time and part-time workers in 2019. The student demographics in the district were as follows: White (55%), Black (18.8%), Hispanic (17%), Asian (4.6%), Multiracial (4.4%), and Native American (0.2%). For the 2017-2018 school year, the district’s school and enrollment figures are displayed in Table 2.

Table 2: *School District Enrollment Demographics*

Schools	Number	Enrollment
Pre-K	Not Available	2,659
Elementary	74	40,154
Elementary/Middle (K-8)	3	2,210
Middle	21	19,752
High	18	28,904
Exceptional	5	688
Charter	19	6,085
Virtual	Not Available	214
Technical College	Not Available	5,172
Adult General Education	Not Available	18,441.

**The District’s Evaluation Model.** The district adopted and began evaluating its teachers and instructional support personnel using the Marzano Teachers Evaluation Model in 2013. At

the time of its adoption, the district had the expectation that the evaluation system would introduce multiple measures of teacher performance that could be paired with professional development, to provide the training and feedback teachers needed to improve their practice and performance. For the 2018-2019 school year, the district adopted the Marzano Focused Teacher Evaluation Model (Carbaugh, Marzano, & Toth, 2017), a revised version of the original model. According to Carbaugh et al. (2017), the revised version is easy to use and less complicated for evaluators compared with the original one. The Marzano Focused Teacher Evaluation Model is divided into four key areas with 23 elements, rated with a five-point proficiency scale. The four key areas include (1) standards-based planning, (2) standards-based instruction, (3) conditions for learning, and (4) professional responsibilities. Each of these areas is weighed differently for the summative classroom observation score. The standards-based planning component accounts for 14%, standards-based instruction (34%), conditions for learning (24%), and professional responsibilities (18%). The five-point performance scale includes: highly effective, effective, developing/needs improvement, and unsatisfactory.

With the Marzano model, the district evaluated all instructional personnel at least once a year. New teachers were, however, observed and evaluated at least two times per year. The evaluation and observation of instructional staff using the Marzano rubric constitutes the instructional practice component of the summative evaluation for the year.

Two other components are the student performance and deliberate practice score components. The deliberate practice component is aimed at helping teachers improve their practice by identifying and participating in relevant professional development activities with the view to increasing student achievement. Teachers were expected to reflect on their practice and professional learning as it relates to their contribution to student growth. To complete the

professional deliberate plan, teachers were required to complete a self-assessment, review schoolwide improvement initiatives, student assessment data, and the previous year's evaluation rating to identify learning goals that will improve student achievement. The student performance component was calculated based the achievement data of students assigned to a teacher over at least a period of three years. The different components were finally weighted for the summative evaluation score as follows: instructional practice (56.7%), student performance data (33.3%), and deliberate practice (10%).

### **Human Subjects Approval**

The University of South Florida (USF) requires the Institutional Review Board (IRB) to review and approve research processes, procedures and instruments to be used in all studies involving human subjects. An application for the study was submitted for review and approval once the study was approved by the dissertation committee. Following several iterations of feedback and revisions, the USF IRB determined that the study would be approved contingent on approval from the school district. The researcher subsequently applied to conduct research in the selected school district. The district granted a preliminary approval within two weeks and that was in turn submitted to the university IRB for final approval. With the district's preliminary approval, the university approved the study (see Appendix 6). The district's final approval was also granted within a few days after final clearance from the university IRB.

### **Procedures**

The survey was administered using a Web-based approach. Web-based survey approaches use internet-based questionnaires or instruments for data collection through online survey platforms like Survey Monkey or Qualtrics (Creswell, 2015). Researchers have identified a number of advantages to using Web-based data collection methods, including their ability to

collect large volumes of data within a short amount of time (Creswell, 2015). They are also inexpensive and flexible, and allow for quicker retrieval of data for analysis (Ary et al., 2010; Check & Schutt, 2012). On the other hand, the response rate for Web-based survey are lower compared to other methods. Samples in Web-based surveys are also limited to participants with Internet access only. Creswell (2015) listed some factors that contribute the problem of low response rates, such as: nonrandom sampling, technological issues, security problems, and problems with junk e-mail.

Once final approvals from the University of South Florida IRB and the school district were granted, the survey instrument was uploaded to Qualtrics. The district was sent the link to preview and certify that it was line with the application submitted for the study. According to Ary et al. (2010), electronic mail surveys are especially appropriate if the researcher is able to obtain the email addresses of all members of a finite population, as in the case of all secondary special education teachers in the selected district. The district subsequently emailed the researcher an Excel file containing the email addresses of all special education teachers in the middle and high schools within the district. The total number of emails contained in the file was 366. The district however cautioned about the possibility that some on the list may have moved to new positions and may not qualify for the study.

Next, participants received an email invitation (Appendix 3) to participate. Using a Google add-on known as Yet-Another-Mail Merge, the email invitations were sent in bulk, as requested by the district's approval letter, but personalized to each participant. There is research to show that surveys addressed to individual participants, not as part of a large mailing list, tend to elicit higher response rates (Ary et al., 2010). According to Creswell (2015), response rate for surveys will vary depending on proper notification, adequate follow-up, interest in the study,

quality of instrument, and incentives. With these factors in mind, Creswell (2015) recommended that researchers use multiple strategies to encourage higher participation and return rates. The email contained a brief background of the study, why they were being invited to participate, and any potential benefits or harm. It further provided a brief description of qualifying criteria to help the potential participants decide if they qualified to complete the survey or not. Finally, the email contained an embedded link to the survey that those who were willing to participate clicked.

To administer a survey, Creswell outlined three-step administration process researchers can follow: (1) mail out the questionnaire/instrument, (2) follow up within two weeks on participants who have not responded, and (3) follow up after additional 2 weeks to remind participants to complete the survey. The actual email containing an invitation to participate in the study, as well as Web-based link, was sent to all 366 potential participants on April 1, 2019. When participants clicked on the link, they were first introduced to the electronic consent form. The consent form (see Appendices 1 and 2) provided detailed background information about the study, any potential benefits and harm, how data would be used and stored, and how participants' confidentiality and privacy would be protected. Finally, it provided the opportunity for participants to give full consent before proceeding to complete the survey.

One week after the initial invitation, a first reminder email was sent out to encourage participation. Two other reminders were sent in the final two weeks of data collection with the final reminder sent on April 22, 2019. The data collection period started on April 1, 2019 and ended on April 26, 2019. The choice of timeline was informed by the IRB approval processes both at the university and school district levels. In December 2018, the researcher applied to conduct the study with a southwest Florida School district that used the Marzano Evaluation Framework to evaluate its teachers. About six weeks later (February, 2019), however, the district

denied the application because they had adopted a new and different evaluation system called the Florida Educators Accomplished Practices (FEAPs). The researcher then applied to the school district where data was collected in February 2019. Final approval was granted in March 2019.

By the end of the data collection period, the researcher had received emails from nine potential participants who indicated that they had changed positions and therefore did not meet the inclusionary criteria set out in invitation emails or the consent form. Hence the total number of qualifying participants amounted to 357. Although the use of incentives is recommended by researchers as one of the ways to encourage participation, the researcher was not able to incentivize participants in the study, in part due to the cost implications emanating from the anticipated large number of participants.

### **Ethical Implications**

Survey studies have been found to pose fewer ethical concerns compared with experimental research (Check & Schutt, 2012). That, notwithstanding, the researcher took appropriate steps to ensure that ethical practices were adhered to and dilemmas addressed. Participants were informed, in the initial invitation to participate, that participation was entirely voluntary and that they could decline participation or opt out at any time without repercussions. To help participants make an informed decision, the informed consent page, included at the start of the survey, stated the motivation for the study as well as any perceived benefits to participants. Two ethical issues that typically arise in survey research relate to issues of confidentiality and anonymity.

**Confidentiality.** According to Check and Schutt (2012), confidentiality is the main ethical concern in survey studies. Researchers must ensure that only authorized personnel have access to the information that can be used to connect respondents to their responses. Researchers



also need to ensure that that information collected is safe and secure and use only numbers, not names, to identify respondents to their responses. The researcher administered the surveys electronically through a trusted and secure website provided by the university, thus limiting access from unauthorized persons. Participants could complete the survey from locations of their choice; in their classrooms and homes, or any other locations they judged safe and private. In addition, the study was not conducted by a team of researchers or with any assistance from other individuals. Following data collection, the researcher had sole custody of data, thus limiting potential information sharing with others not involved in the research (Creswell, 2015). Finally, data was stored on a secure cloud-based storage site procured by the university to ensure the privacy of data.

**Anonymity.** Ensuring anonymity requires that researchers do not collect identifying information that can be used to link participants to their responses (Check & Schutt, 2012). To protect the anonymity of participants, personally identifying information such as names, social security numbers, school site, internet protocol (IP) addresses, and other personal details were not collected (Woolf, 2018). Participants who met the inclusionary criteria were sent a link through which they could anonymously complete the survey. In analyzing the data collected, caution was also taken not to report subsets of findings that could potentially reveal the identity of specific participants (Creswell, 2015).

### **Data Analysis**

According to Creswell (2015), data analysis in survey research consists of the following steps: note the response rate, check for response bias, conduct descriptive analysis of all items, and answer research questions. The survey administration closed on April 26, 2019. After the administration deadline, the researcher checked the response rate and reviewed individual

responses for potential response bias or incomplete datasets to prepare the raw data for analysis. Out of the 357 potential participants, 113 responded by completing the survey. The responses came from all (except one, with two participants), 21 middle, 18 high, and five special schools. All 113 responses could not, however, be included in the analysis. The researcher discovered that 17 responses were completely blank, an indication that the respondents simply clicked through the survey items without responding to them. Those responses were deleted. Following the initial review, 96 responses representing 26.9% response rate, were included in the formal analysis of data. The analysis was done into two phases.

In the first phase, data were exported from the internet hosting site (Qualtrics), first into a Microsoft Excel spreadsheet. Microsoft Excel was used to clean and code all the data for analysis. In the second phase, the coded data generated from Excel was uploaded into the Statistical Package for the Social Sciences (SPSS) program to analyze trends in the data. SPSS was first used to calculate mean responses for all individual responses and for response categories. To answer the research questions guiding the study, responses were clustered into three main themes: *clarity*, *practicality*, and *cost*. The first cluster consisting of three items was coded for clarity. Items dealing with *clarity* measured participants' clarity or understanding of the evaluation framework and its subcomponents such as classroom observation rubrics and student growth metrics. *Clarity* items were analyzed for research question one. The second cluster, named as *practicality* was composed of 13 items that were further divided into four measures of evaluator perceptions, feedback, student growth, and professional development. The 13 *practicality* items were used analyzed for research questions two and three. The last cluster of three items, referred to as *cost*, was analyzed for the overall cost perception associated with the evaluation system as a whole. Job category responses were collapsed into three groups of special

education settings: general education, self-contained, and special school. The three groups had the following number of participants in the final number of responses included in the data analysis: general education (40), self-contained (45), special school (11).

Data were analyzed using both descriptive and inferential statistics to organize and summarize the data for easier comprehension. To account for missing data, the researcher used the average responses for item clusters, thus allowing for respondents who skipped some questions to still be counted. For example, respondents who answered two out of the three items clustered for clarity were included in the analysis. Descriptive statistics included frequency distributions, and measures of central tendency like mean, mode, and median. The descriptive statistics provided a general understanding of trends in the data. The findings from the descriptive analysis were used to answer the first three research questions. Apart from descriptive statistics, inferential statistics such as independent *t*-tests and analysis of variances (ANOVA) were also computed to study patterns and draw conclusions from participants' responses in relation to the research questions. Inferential statistics proved especially useful in analyzing variations in perceptions across the demographic groups of gender, job categories, years of experience. It also helped to establish whether participants were drawn from the same population of special education teachers.

The final item on the instrument was the only open-ended item. It asked participants to share any other comments they had about the evaluation framework used in their district. More than half of the participants provided comments that were analyzed for themes using the implementation factors identified in the conceptual framework. From the analysis, most of the comments fell under *practicality* and its subcategories of feedback received, professional

development, and responsiveness to job descriptions. A selection of the comments was woven into the discussions of the major findings of the study and presented in Chapter Five.

### **Summary**

This chapter has specified the research methods used for data collection and analysis. Quantitative approaches, through census survey design was adopted, due to the usefulness of surveys in describing trends, determining people's opinions regarding policy issues, and identifying important beliefs and attitudes of individuals. Data were collected electronically using a questionnaire and analyzed with descriptive and inferential statistics. The researcher made conscious efforts to ensure compliance with ethical practices in survey research and with university and district IRB requirements.

## **CHAPTER FOUR: RESULTS**

### **Overview**

The purpose of the study was to explore special education teacher perceptions regarding the teacher evaluation system that is used to evaluate their performance and effectiveness. Specifically, the study sought to find out whether special educators viewed the evaluation tool as responsive to their job descriptions, pedagogical and practice needs, as well professional development aspirations. Using a quantitative research approach, the researcher adopted a survey instrument and surveyed all special education teachers at the secondary level (middle and high school) in a southwest Florida school district that uses the Marzano Evaluation Framework to evaluate all teachers, including special educators. A census survey approach that allows researchers to study an entire population when it is small and can be identified without much effort, was used to for this study. All special education teachers at the middle and high school levels, as identified by the school district, received an invitation to complete a survey for the study. The following research questions guided the exploration of special educator perceptions pertaining to teacher evaluation systems.

- How do special education teachers rate their understanding of the teacher evaluation system used in their district?

- What are the perceptions of special education teachers about teacher evaluation systems with regard to their multiple roles, responsibilities, and contexts?
- What are the perceptions of special education teachers about evaluations systems with regard to the ability of these systems to address professional development, practice, and pedagogical needs?
- Do special education teachers' perceptions of teacher evaluation systems vary by:
  - a) gender
  - b) job categories
  - c) years of experience.

The data analysis was guided by the theoretical framework undergirding the study. The theoretical framework is based on the theories of implementation, as advanced by Doyle and Ponder (1977) and Fullan (2001), who posit that teacher perceptions of policies and their implementation are mediated by clarity, practicality and cost perceptions. The data were thus categorized and analyzed through the lens of those factors of implementation. The analysis was performed through SPSS for both descriptive and inferential statistics. The descriptive statistics presented in this chapter include number of respondents for each item, the means, standard deviations, skewness, kurtosis, and minimum and maximum scores. Reliability ( $r$ ), using Cronbach's alpha, is also presented for clusters of items that were analyzed together to answer the research questions. For inferential statistics, the results include two-tailed independent  $t$ -tests and one-way between subjects analysis of variance (ANOVA) that were conducted to compare groups of participants in the study.

This chapter presents the results of the analysis. The presentation is done by research questions. Descriptive statistics for all research questions are presented first followed by

inferential results for applicable research questions. The results suggest that special education teachers do not have a high level of clarity of the evaluation framework used to rate their performance. Participants rated themselves as having low understanding about how the different components of the evaluation process are combined to generate a summative score of their effectiveness. Further, the special educators surveyed did not view the existing evaluation frameworks as practical in distinguishing their differing roles and responsibilities as well and their practice and professional development needs. The results also suggest that special educators perceive a high cost in terms of the emotional toll and anxiety associated with evaluation systems. Detailed discussion and implications of the results will be presented in Chapter 5.

## **Results**

Research Question 1: *How do special education teachers rate their understanding of the teacher evaluation system used in their district?*

A subgroup of items classified as *clarity* that sought participants' understanding of the evaluation system was analyzed to answer the first research question. The items asked participants to rate their understanding of the evaluation framework, how different assessments are combined to generate student growth component of summative evaluation scores, and how observations and student growth are combined to generate a summative performance score. Using a Likert scale of one to five (5 = excellent, 4 = very good, 3 = good, 2 = fair, 1 = poor), participants indicated their understanding of the evaluation framework and its components. Descriptive statistics of the findings are displayed in Table 3.

Table 3: *Perceptions of Clarity*

Item	N	Min.	Max.	<i>M</i>	<i>SD</i>
The observation rubric or framework used to rate special education teacher professional practice.	96	1	5	2.62	1.20
How different assessments are combined to create student growth measure.	96	1	5	2.49	1.17
How different observations and student growth measures are combined to determine a summative performance score	95	1	5	2.47	1.16

Table 3 shows that participants had a fair understanding of the evaluation rubric that is used to rate their practice ( $N = 96$ ,  $M = 2.62$ ,  $SD = 1.20$ ). In terms of how assessments are combined to generate student growth scores, participants understanding was similar ( $N = 96$ ,  $M = 2.49$ ,  $SD = 2.49$ ). The results further show that participants ( $N = 95$ ) did not have a high level of understanding regarding how student growth and classroom observations are combined to determine a composite performance score ( $M = 2.47$ ,  $SD = 1.16$ ).

All the *clarity* items were also combined and analyzed for an overall score for participants' perception regarding their clarity of the evaluation framework. The findings, which indicated that participants ( $N = 96$ ) had an overall average (fair) understanding ( $M = 2.53$ ,  $SD = 1.06$ , Skewness = 0.44, Kurtosis = -0.58) are shown in Table 4.

Table 4: *Participant's Overall Perceptions of Clarity*

<i>N</i>	<i>M</i>	Min.	Max.	<i>SD</i>	Skewness	Kurtosis	Cronbach's alpha*
96	2.53	1	5	1.06	0.44	-0.58	.89

\* Three-item subscale



Research Question 2: *What are the perceptions of special education teachers about teacher evaluation systems with regard to their multiple roles, responsibilities, and contexts?*

To answer the second research question, another subgroup of items categorized as *practicality* was analyzed. Practicality items solicited participants perceptions about the evaluation system relative to their evaluator, roles, responsibilities, and contexts, pedagogical and practice needs, as well as professional development considerations. There were 17 *practicality* items that were further grouped into four measures: evaluator, feedback, student growth, and professional development perceptions. Three of the four measures, comprising 13 out of the 17 items, were analyzed to answer research question 2. The remaining four items were analyzed to answer the third research question.

Perceptions about the first *practicality* measure, the evaluator, dealt with how participants viewed the evaluator who assessed them relative to the evaluator’s ability to assess their instruction, roles, responsibilities, and practice. Participants rated on a Likert scale of one to four (4 = to a great extent, 3 = to some extent, 2 = a little, 1 = not at all) their perceptions regarding their evaluator’s understanding of the participants’ classroom practice, including strengths and weaknesses, and whether or not the evaluator did so in a fair and unbiased manner. The results showed that participants perceived their evaluators in positive ways. Table 5 depicts the results.

Table 5: *Practicality: Perceptions about Evaluator*

Item	<i>N</i>	Min.	Max.	<i>M</i>	<i>SD</i>
Is able to accurately assess my instruction, roles, responsibilities, and practice.	96	1	4	3.08	0.84
Understands my strengths and weaknesses as a special educator.	96	1	4	3.08	0.90
Is fair and unbiased.	96	1	4	3.31	0.87
Understands what is going on in my classroom	93	1	4	3.06	0.94

The findings in Table 5 show that participants had favorable perceptions regarding their evaluators' ability to evaluate their instruction, roles, and responsibilities ( $M = 3.08$ ), understand their strengths and weaknesses ( $M = 3.08$ ), and understand what happens in a special education classroom ( $M = 3.06$ ). Participants' views on the fairness and unbiasedness of their evaluators had the highest mean ( $M = 3.31$ ).

The overall perception about evaluators is displayed in Table 6. Participants ( $N = 96$ ) agreed to some extent ( $M = 3.14$ ,  $SD = 0.77$ , Skewness =  $-0.97$ , Kurtosis =  $0.55$ ) that their evaluators were capable of assessing their classroom practice, and roles, responsibilities, and practice as special educators in a fair and unbiased manner.

Table 6: *Practicality: Overall Perceptions about Evaluator*

<i>N</i>	<i>M</i>	Min.	Max.	SD	Skewness	Kurtosis	Cronbach's alpha*
96	3.14	1	4	0.77	-0.97	0.55	.88

\*Four-item subscale

On a scale of one to five (5 = strongly agree, 4 = agree, 3 = neither agree nor disagree, 2 = disagree, 1 = strongly disagree), participants rated their views on the second practicality measure – feedback. The six items on this measure sought perceptions about usefulness of feedback and the appropriateness of the evaluation scales in terms of their roles, responsibilities, and practices. The means and standard deviations together with the minimums and maximums of all responses are shown in Table 7.

Table 7: *Practicality: Perceptions about Feedback*

Item	<i>N</i>	Min.	Max.	<i>M</i>	<i>SD</i>
The rating scale used to assess my roles, responsibilities, and practices as a special education teacher are appropriate.	95	1	5	2.42	1.15
I am given useful feedback by the evaluator.	95	1	5	3.63	0.90
The feedback I received this year identified specific areas of my instruction, roles, responsibilities, and practice that could be improved.	95	1	5	3.47	0.99
I have used the feedback I received this year to improve my teaching, roles, responsibilities, and practice.	93	1	5	3.47	1.06
The feedback I received this year included guidance on how to make improvements to my instruction and special education practice.	94	1	5	3.10	1.10
I would recommend a different evaluation rubric for special education teachers compared to general education teachers.	94	1	5	4.54	0.82

The results displayed in Table 7 suggest that participants disagreed that using the existing evaluation to rate their roles, responsibilities, and practice as special education teachers was appropriate. The mean score for that item was lower ( $M = 2.42$ ) on a scale of one to five. Participants were also not sure of the usefulness of the feedback received from their evaluators as the mean score of 3.63 (within the range of neither agree nor disagree and agree) indicate when they were asked to indicate their agreement or disagreement relative to the usefulness of feedback received. As to whether participants believed that the feedback they received identified areas of improvement or they had used the feedback received to improve their practice and pedagogy, or the mean scores for both items were the same ( $M = 3.47$ ), indicating a not sure response. The final item on feedback asked participants if they would recommend a different evaluation rubric for special educators as opposed to the current system of using rubrics

validated for general education teachers. The strong mean of 4.54 ( $SD = 0.82$ ) indicated strong support to have a separate evaluation framework for special education teachers.

Altogether, participants were not sure of the practicality of the feedback they received and the appropriateness of the evaluation rubric in assessing their roles, responsibilities, and practices as special educators. The overall mean score for the feedback measure of practicality was 3.44 ( $SD = 0.71$ ) on a scale of one to five. It suggests that participants neither agreed on the appropriateness of the evaluation rubric nor the usefulness of the feedback received. Table 8 shows the overall feedback indicator of practicality.

Table 8: *Practicality: Overall Perceptions about Feedback*

<i>N</i>	<i>M</i>	Min.	Max.	SD	Skewness	Kurtosis	Cronbach's alpha*
95	3.44	1	5	0.71	-0.44	0.06	.77

\*Six-item subscale

The third measure used to answer research question 2 was the practicality measure referred to as student growth. The measure consisted of three items that asked participants to indicate their agreement or disagreement using a scale of one to five (5 = strongly agree, 4 = agree, 3 = neither agree nor disagree, 2 = disagree, 1 = strongly disagree) relative to questions about the student growth component of the evaluation system. The means scores suggest that most participants have made changes to their teaching and practice with the view to improving their students' performance on assessments that are used in the student performance component of the evaluation system. Although participants agreed ( $M = 3.51$ ,  $SD = 0.825$ ) to have made changes to their teaching practice ( $M = 3.51$ ) to improve their students' scores, they disagreed with the statement that the student growth measure of the evaluation framework is a fair representation of their students' learning ( $M = 2.05$ ,  $SD = 1.06$ ). They further disagreed with

using the information generated from those assessments to inform their professional development choices ( $M = 2.58$ ,  $SD = 1.14$ ). Table 9 shows the scores of student growth practicality measure.

Table 9: *Practicality: Perceptions about Student Growth*

Item	<i>N</i>	Min.	Max.	<i>M</i>	<i>SD</i>
The measures of student growth are a fair representation of my students' learning.	93	1	5	2.05	1.06
The information I get from those assessments will inform my professional development choices.	93	1	5	2.58	1.14
I have made changes in my teaching and practice in order to improve my students' scores on these assessments	93	1	5	3.51	1.13

The average scores for the student growth measure of practicality revealed that participants disagreed with using student growth measures to evaluate their performance. The mean score ( $M = 2.73$ ) ranged between disagree and neither agree nor disagree. Although they agreed to using the information from the assessments for student growth to make changes to their practice, they appeared to have done so only for the purpose of improving their students' performance on those assessments, which could ultimately improve their own score student growth component of the final evaluation rating. Table 10 shows that average scores for the practicality measure of student growth.

Table 10: *Practicality: Overall Perceptions about Student Growth*

<i>N</i>	<i>M</i>	Min.	Max.	<i>SD</i>	Skewness	Kurtosis	Cronbach's alpha*
93	2.71	1	5	0.91	-0.32	-0.78	.77

\*Three-item subscale

Research Question 3: *What are the perceptions of special education teachers about evaluations systems with regard to the ability of these systems to address professional development, practice, and pedagogical needs?*

The final practicality measure classified as professional development was analyzed for research question 3. The four items under this measure asked participants to rate the extent to which they agreed (5 = strongly agree, 4 = agree, 3 = neither agree nor disagree, 2 = disagree, 1 = strongly disagree) with statements about using the evaluation process to inform their professional development choices and practice. The findings are shown in Table 11.

Table 11: *Practicality: Perceptions about Professional Development*

Item	<i>N</i>	Min.	Max.	<i>M</i>	<i>SD</i>
The observation process encouraged me to reflect on my teaching practice.	95	1	5	3.45	1.02
My observation ratings will guide my future professional development choices.	95	1	5	3.06	1.00
My observation results will guide my future professional development activities.	95	1	5	3.02	0.98
I have made changes in my teaching and practice as a result of the observation process,	94	1	5	3.25	1.05

Results showed that participants' responses to being encouraged to reflect on their teaching practice as a result of the evaluation process fell within the range of neither agree nor disagree and agree ( $M = 3.45$ ). Participants were also not sure about using observation data to guide their professional development choices ( $M = 3.06$ ), influence future professional development activities ( $M = 3.02$ ), and make changes in teaching practice as a result of the observation process ( $M = 3.24$ ).

Based on the overall set of responses for the professional development measure of practicality, participants were unsure with regard to the ability of evaluation system to address their professional development and practice needs. The overall mean score of the professional development measure was 3.2 ( $SD = 0.89$ ) representing neither agree nor disagree. The overall scores are presented in Table 12.

Table 12: *Practicality: Overall Perceptions about Professional Development*

<i>N</i>	<i>M</i>	Min.	Max.	SD	Skewness	Kurtosis	Cronbach's alpha*
95	3.2	1	5	0.89	-0.25	-0.19	.89

\*Four-item subscale

**Perceptions about Cost:** The final set of items on the instrument was labelled for Cost perceptions to ascertain the impact of the evaluation system on participants. Questions about cost dealt with participants' perceptions about the ease of the evaluation process and any potential benefits or returns from the entire process. There were three questions associated with cost. The first two asked participants to indicate their agreement or otherwise on a scale of one to five (5 = strongly agree, 4 = agree, 3 = neither agree nor disagree, 2 = disagree, 1 = strongly disagree) about the level of anxiety and effort associated with the evaluation process. The results show that participants were of the view that the evaluation process has increased their level of stress and

anxiety ( $M = 3.96$ ). As a result, participants agreed that the evaluation process was not worth the effort. This is indicated by the high mean that fell within the range of agree and strongly agree ( $M = 4.08$ ). Table 13 shows the outcome of the cost analysis.

Table 13: *Cost Perceptions*

Item	<i>N</i>	Min.	Max.	<i>M</i>	<i>SD</i>
The evaluation process has increased my level of stress and anxiety.	95	1	5	3.96	1.15
Overall, the evaluation process takes more effort than the results are worth.	95	1	5	4.08	1.14
Overall, what kind of effect do you think the existing evaluation has had on your effectiveness as a special education teacher?	95	1	5	2.92	0.91

The final cost question in Table 13 asked participants provide an overall rating for their existing evaluation system in term of its effect on their effectiveness as special educators. Using a scale of one to five (5 = strongly negative effect, 4 = positive effect, 3 = no effect, 2 = negative effect, 1 = strongly negative effect), responses indicate that participants perceived the evaluation system as having negative to no effect ( $M = 2.92$ ) on their effectiveness as special education teachers.

The overall mean ( $M = 4.02$ ) of all cost items shows participants perceived a high cost in terms of the stress and anxiety associated with the evaluation process. Table 14 displays that overall cost perception scores.



Table 14: *Overall Cost Perceptions*

<i>N</i>	<i>M</i>	Min.	Max.	Skewness	Kurtosis	Cronbach's alpha*
95	4.02	1	5	-1.10	1.03	.70

\*Three-item subscale

Research Question 4: *Do special education teachers' perceptions of teacher evaluation systems vary by:*

- a. gender*
- b. job categories*
- c. years of experience.*

To answer the final research questions, inferential statistics were used to compare perceptions across groups. Comparisons were done using means of participants' gender, years of experience, and job categories to establish if there were statistically significant differences among the participant groups with respect to their perceptions of the evaluation system. Specifically, two tests were conducted: independent *t* tests by gender and analysis of variances (ANOVA) for job categories and years of experience on the job. Independent samples *t*-tests were used for gender because it had only two groups (male and female). The other variables (years of experience and job categories) had more than two groups. Thus, the researcher deemed that ANOVA would be appropriate to test between group variability instead of independent samples *t*-tests.

Independent *t*-tests. The independent *t*-test compared gender with overall perceptions about clarity and cost, as well as all the sub-categories of practicality of the evaluation framework. Findings of each measure are presented, starting with a comparison of gender and the clarity. It is followed by the results of comparisons between gender and practicality and cost

perceptions. The results show that there are no statistically significant differences between male and female and special education teachers' perceptions regarding evaluation systems.

Table 15 exhibits the group statistics for the three perception indicators while Table 16 shows the independent samples t-test scores.

Table 15: *T-test: Group Statistics of Clarity, Practicality, and Cost by Gender*

	Gender	<i>N</i>	<i>M</i>	<i>SD</i>	<i>t</i> -statistic
Clarity	Male	22	2.43	0.78	-0.46 ns
	Female	72	2.55	1.14	
Practicality _ Evaluator	Male	22	3.31	0.64	1.31 ns
	Female	72	3.07	0.80	
Practicality - Feedback	Male	22	3.58	0.60	1.03 ns
	Female	72	3.40	0.75	
Practicality – Student Growth	Male	21	2.80	0.93	0.52 ns
	Female	71	2.68	0.91	
Practicality – Professional Development	Male	22	3.20	0.90	0.06 ns
	Female	72	3.19	0.89	
Cost	Male	22	4.00	1.02	-0.17 ns
	Female	72	4.04	1.00	

Note: ns = not statistically significant ( $p > .05$ )

Table 15 shows the results of the independent samples t-test that was conducted to compare gender and clarity, practicality, and cost perceptions. The results found no statistically significant differences between males ( $M = 2.95$ ,  $SD = 0.89$ ) and females ( $M = 2.89$ ,  $SD = 0.91$ ) on clarity perceptions;  $t(92) = -0.46$ . The results suggest that participants' gender has no effect relative to their understanding (clarity) of the observation system ( $p > .005$ ). In terms of gender and perceptions regarding practicality, no differences were observed between males and females for any of the practicality measures. The *t*-statistic for each of the four practicality measures was

not significant: evaluator,  $t(92) = 1.31$ ; feedback,  $t(92) = 1.03$ ; student growth,  $t(90) = 0.52$ ; and professional development,  $t(92) = 0.06$  ( $p > .005$ ). The results indicate that gender is not associated with special educators' perceptions about the practicality of the evaluation system. The test for the relationship between gender and cost also had a similar outcome. At the  $p < .05$  level, the independent t-test showed no differences between males ( $M = 4.00$ ,  $SD = 1.02$ ) and females ( $M = 4.04$ ,  $SD = 1.00$ ) cost perceptions;  $t(92) = -0.17$  ( $p > .005$ ).

Table 16: *One-Way ANOVA of Clarity, Practicality, and Cost Perceptions by Years of Experience*

		Sum of Squares	df	Mean Square	F	Sig.
Clarity	Between Groups	.998	2	.499	.439	.646
	Within Groups	105.779	93	1.137		
	Total	106.777	95			
Practicality - Evaluator	Between Groups	.576	2	.288	.486	.616
	Within Groups	55.080	93	.592		
	Total	35.675	95			
Practicality - Feedback	Between Groups	.247	2	.123	.240	.787
	Within Groups	47.360	92	.515		
	Total	47.606	94			
Practicality – Student Growth	Between Groups	2.319	2	1.160	1.397	.253
	Within Groups	74.701	90	.830		
	Total	77.020	92			
Practicality – Professional Development	Between Groups	.868	2	.434	.545	.581
	Within Groups	73.207	92	.796		
	Total	74.075	94			
Cost	Between Groups	2.462	2	1.231	1.231	.297
	Within Groups	91.996	92	1.000		
	Total	94.458	94			

**One-way ANOVA tests.** One-way between subjects ANOVA tests were conducted to compare perceptions by years of experience and job descriptions. The first one-way between subjects ANOVA was conducted to compare the effect of years of experience on participants'

perceptions of clarity, practicality (evaluator, feedback, student growth, and professional development), and cost perceptions of the evaluation system. Table 16 shows the results of years of experience by perceptions of clarity, practicality (evaluator, feedback, student growth, and professional development), and cost.

The one-way between subjects ANOVA conducted to compare participants' years of experience and perceptions of clarity indicated that there was no statistically significant association between years of experience and clarity at the  $p < .05$  level;  $F(2, 93) = 0.44, p = .646$ . The ANOVA comparing years of experience and perceptions of the four practicality measures also showed no significant effect at the  $p < .05$  alpha level: evaluator,  $F(2, 93) = 0.486, p = .616$ ; feedback,  $F(2, 92) = 0.240, p = .787$ ; student growth,  $F(2, 90) = 1.397, p = .253$ ; and professional development,  $F(2, 92) = 0.545, p = .581$ . The final ANOVA conducted to test the effect to years of experience and perceptions of cost also revealed no significant differences;  $F(2, 92) = 1.23, p = 0.279 (p > .005)$ . Altogether, the results suggest that that years of experience on the job has no impact on how special education teachers view the evaluation system that is used to evaluate their performance.

Table 17: *One-Way ANOVA of Clarity, Practicality, and Cost Perceptions by Job Categories*

		Sum of Squares	df	Mean Square	F	Sig.
Clarity	Between Groups	1.532	2	.766	.677	.511
	Within Groups	105.245	93	1.132		
	Total	106.777	95			
Practicality - Evaluator	Between Groups	1.323	2	.662	1.133	.327
	Within Groups	54.333	93	.584		
	Total	55.656	95			
Practicality - Feedback	Between Groups	.889	2	.444	.875	.420
	Within Groups	46.718	92	.508		
	Total	47.606	94			
Practicality – Student Growth	Between Groups	.835	2	.418	.493	.612
	Within Groups	76.185	90	.846		
	Total	77.020	92			
Practicality – Professional Development	Between Groups	3.711	2	1.855	2.426	.094
	Within Groups	70.364	92	.765		
	Total	74.075	94			
Cost	Between Groups	1.096	2	.548	.540	.585
	Within Groups	93.362	92	1.015		
	Total	94.458	94			

A second set of one-way ANOVAs was conducted to establish if perception varied by the job categories of participants. A total of three comparisons were conducted. The findings of the comparisons between the job categories and perceptions of clarity, practicality, and cost are presented in Table 17. As shown in Table 17, the one-way ANOVA conducted to verify if the participants' perceptions of clarity varied by their job categories shows there is no statistically-significant difference at the alpha level of  $p < .05$ ;  $F(2,93) = 0.68$ ,  $p = 0.511$ . The comparison of

job categories and the practicality measures revealed a similar outcome of no difference as follows: evaluator,  $F(2,93) = 1.133, p = .327$ ; feedback,  $F(2, 92) = 0.875, p = .420$ ; student growth,  $F(2, 90) = 0.493, p = .612$ ; and professional development,  $F(2, 92) = 2.426, p = .094$ . The third comparison was between job description and participants' views about cost. The results established that job description or category had no effect on participants' cost perceptions;  $F(2, 92) = 0.54, p = 0.585$  ( $p > .005$ ). The combined results suggest that special education teachers' perceptions about the existing evaluation systems are not mediated by their job categories or descriptions. Special education teachers in different settings and contexts have similar opinions in terms of the evaluation frameworks used to evaluate them.

### **Summary of Results**

This chapter presented the results from the data obtained from the survey administered for the study. The data were analyzed using descriptive and inferential statistics and generated information about trends in respondents' perceptions with regard to the evaluation system used in their district. The theoretical framework underpinning the study was also used to organize the data for easy interpretation and alignment with the research questions.

The findings reveal that respondents did not have a high level of understanding about the evaluation framework used in their district. Their mean level of understanding was only rated as fair. Participants' mean levels of understanding regarding how different assessments are combined to generate student growth measures and how student growth measures are combined with observations to determine a summative performance score was even lower.

The findings also revealed participants' perceptions about the practicality of the evaluation system in relation to their multiple roles, responsibilities, classroom contexts, and professional development needs. The perceptions about practicality were analyzed across four

different measures: evaluator, feedback, student growth, and professional development. The findings suggest that participants had positive views of their evaluators judging by the high mean score obtained for the evaluator measure of practicality. The other indicators of practicality, however, showed that participants did not agree that the evaluation framework was appropriate for their roles and responsibilities. Although participants agreed that they had used the information and feedback from the evaluations to improve instruction, they only did so for the purposes of improving their students' performance on the assessments used to generate the student growth component of their performance score. The findings also indicate that participants were opposed to using student growth measures as part of the evaluation process. One notable finding from the results is the strong agreement from participants that a separate evaluation framework should be designed for special education teachers.

Another pertinent finding related to participants' views regarding the perceived anxiety and stress associated with the evaluation system. The findings showed that participants felt the evaluation process exacts a heavy emotional toll and the effort it takes are not worth it. When finally asked to estimate the overall effect that evaluation process has had on their effectiveness as special education teachers, half of the participants (67.7%) responded that it had no effect, with another 17.71% reporting that the system had a negative effect on their effectiveness.

The findings from the inferential analysis indicate that participants' expressed perception did not differ across groups of special education teachers. The outcome of the t-test that compared gender and perceptions showed that there were no significant differences between male and female participants regarding viewpoints on the clarity, practicality, and cost associated with the evaluation system used in their district. The one-way ANOVA comparisons also showed

that participants' perceptions about their evaluation framework did not differ regardless of their job categories or their years of experience of the job.

The findings will be discussed in tandem with the literature on the subject of special education teacher evaluation in the next chapter. Implications for research and practice as well as recommendations will also be presented.



## **CHAPTER FIVE: DISCUSSION**

### **Overview**

This study was informed by three crucial issues pertaining to measuring special education teacher performance through the use of existing teacher evaluation models and systems. The issues included: (1) ongoing discussion regarding the challenges associated with evaluating special education teacher effectiveness, (2) the lack of research documenting special education teacher perceptions in terms of how they have fared under those systems, and (3) the enduring issue of special education teacher attrition. The purpose of the study was to examine the special education teacher perceptions related to these three issues. Specifically, the study explored how special education teachers perceived the existing evaluation systems in terms of the ability of the systems to distinguish their multiple and differing roles, responsibilities, contexts, and address their pedagogy, practice, and professional needs.

To conceptualize special education teacher perceptions, the study drew on the theories of Doyle and Ponder (1977) and Fullan (2001) regarding the factors that impact teachers' perceptions relative to the implementation of policies, such as a teacher evaluation system. The researcher was interested in finding out how special education teachers perceived the evaluation systems through the lens of the factors of policy implementation identified by Doyle and Ponder (1977) and Fullan (2001).

The study used quantitative methodologies, employing a census survey approach to elicit the participants' perceptions. An invitation to complete an online survey was sent to all 357 secondary-level (middle and high school) special education teachers in a southwestern Florida school district that used the revised Marzano Focused Teacher Evaluation Model to evaluate its teachers, including special education teachers. A total of 113 participants responded to the online survey with 96, representing a 26.8% response rate, were included in the final analysis. SPSS was used to conduct both descriptive and inferential analysis to answer the research questions guiding the study. The findings of the analysis were presented in the preceding chapter.

The purpose of this chapter is to discuss the findings from the data analysis, within the context of the literature and the framework guiding the study, and draw conclusions regarding the perceptions of special education teachers about existing teacher evaluation systems. Implications and recommendation for practice and future research are also discussed along with limitations of the study.

### **Findings**

Teacher evaluation systems have been adopted by several state and local education agencies for the purpose of assessing teacher effectiveness and contribution to student growth. These evaluation systems have had implications for teacher retention and remuneration (Darling-Hammond, 2013). There have been ongoing discussions about the challenges associated with using existing evaluation systems to assess special education teachers in view of their unique job responsibilities and instructional contexts. Missing from the discussions are the contributions and insights from special education teachers who are impacted by the teacher evaluation policies. This study was aimed at exploring how special education teachers perceive the existing systems specifically as it relates to their: (1) understanding of the evaluation frameworks and

components, (2) perceptions about the ability of the frameworks to address their professional development, pedagogical, and practice needs, and (3) views pertaining to the responsiveness of the systems with regard to their roles, responsibilities, and contexts. The theories of implementation by Doyle and Ponder (1977) and Fullan (2001) that advanced factors that influence teacher perceptions of policies and their implementation, provided a window through which to make sense of the perceptions expressed by participants. Guided by the methods of Jiang et al. (2015), the study combined the factors to guide the interpretation of special education teacher perceptions. The researcher assumed that the perceptions about factors of *practicality*, *clarity*, and *cost* would inform participants' views of existing evaluation system. The research questions addressed in the findings were:

- How do special education teachers rate their understanding of the teacher evaluation system used in their district?
- What are the perceptions of special education teachers about teacher evaluation systems with regard to their multiple roles, responsibilities, and contexts?
- What are the perceptions of special education teachers about evaluations systems with regard to the ability of these systems to address professional development, practice, and pedagogical needs?
- Do special education teachers' perceptions of teacher evaluation systems vary by:
  - a) gender
  - b) job categories, and
  - c) years of experience.

The findings for each of the research questions are presented first, followed by a detailed discussion.

**Findings for Research Question 1:** *How do special education teachers rate their understanding of the teacher evaluation system used in their district?*

The first research question about special education teachers' understanding of the evaluation system aligned with the implementation factor of *clarity* in the theoretical framework. Respondents rated their understanding of the teacher evaluation system used in their district on a 5-point scale (5 = excellent, 4 = very good, 3 = good, 2 = fair, 1 = poor). The findings from the set of items under this research question suggest that participants only had a fair understanding or clarity of the evaluation framework ( $M = 2.62$ ). Their understanding of how different the student growth component of the evaluation is generated was slightly lower ( $M = 2.49$ ) although it was still within the fair to good range. Participants understanding of how the different components of the evaluation framework; in this case classroom observations, deliberate practice plans, and student performance are combined to generate summative performance scores was also within the fair range ( $M = 2.47$ ). The overall clarity or understanding mean score ( $M = 2.53$ ) indicates that participants did not have a deeper understanding of the evaluation framework and how it is used to rate their effectiveness.

**Findings for Research Question 2:** *What are the perceptions of special education teachers about teacher evaluation systems with regard to their multiple roles, responsibilities, and contexts?*

For this research question, 13 items under the umbrella of *practicality* were further categorized into four areas of evaluator perceptions, feedback received, and student growth perceptions. The findings from these items were used to discern participants' viewpoints about the evaluation system relative to their multiple roles, responsibilities, and contexts. The findings indicate that participants had slightly positive perceptions about the evaluators' ability to

evaluate their roles, responsibilities, and contexts ( $M = 3.08$ ). They also felt that their evaluators were somewhat fair in their observation ratings ( $M = 3.31$ ). The overall evaluator perception mean was 3.14 ( $M = 3.14, SD = 0.77$ ).

While participants had slightly positive perceptions about their evaluators, the same could not be said of their perceptions regarding the responsiveness of the evaluation system to their roles, responsibilities, and contexts. The findings from the category of feedback received, revealed that participants disagreed with the appropriateness of using the existing rubrics to evaluate their roles, responsibilities, classroom contexts ( $M = 2.42$ ). They were also not sure of the usefulness of the feedback received ( $M = 3.63$ ). It was not surprising, therefore, that participants could not agree or disagree about using the feedback received to improve their teaching, roles, responsibilities, or practice ( $M = 3.47$ ). One notable finding was participants' strong agreement ( $M = 4.54, SD = 0.82$ ) with the recommendation for a different evaluation rubric for special education teachers.

The final category of items for research question 2 dealt with perceptions about the student growth component of the evaluation process. The findings show that participants were not in favor of using student performance indicators to rate special education teachers. On a scale of 1 to five, participants were of the view that measures of student growth used in their district was not representative of their students' learning and abilities ( $M = 2.05$ ). Participants did not use the information generated from assessments used for student performance indicators ( $M = 2.58$ ). Instead, they made some changes in their teaching and practice ( $M = 3.51$ ) but only with the intent of improving their students' scores of the assessments. The findings from all the categories suggest that participants perceived the evaluation system was not appropriate for rating special education teacher performance relative to their roles, responsibilities, and contexts.

**Findings for Research Question 3:** *What are the perceptions of special education teachers about evaluations systems with regard to the ability of these systems to address professional development, practice, and pedagogical needs?*

A set of three items, also classified as *practicality* were analyzed to answer this research question. The responses to the items were on a scale of 1 to 5 (5 = strongly agree, 4 = agree 3 = neither agree nor disagree, 2 = disagree, 1 = strongly disagree). Participants could not fully agree that the observation process had encouraged them to reflect on their teaching practice ( $M = 3.45$ ) or made changes to their teaching and practice due to the observation process ( $M = 3.25$ ). In terms of their process developments planning and activities, participants were not sure about using the observation feedback to guide their future professional development choices ( $M = 3.06$ ) and activities ( $M = 3.02$ ). The overall findings suggest that participants were not sure regarding the ability of the evaluation system to address their professional development, practice, and pedagogical needs. The overall mean score was 3.2 ( $M = 3.2$ , within neither agree nor disagree).

**Cost Perceptions:** A group of items were analyzed to determine the overall impact of the evaluation system on participants relative to the expressed perceptions. Participants expressed that evaluation process is associated with increased levels of frustration and anxiety. Participants agreed that the evaluation process has increased their levels of stress and anxiety ( $M = 3.96$ , out of a maximum of 5). There was a stronger perception that the results from the evaluation were not worth all the effort it takes ( $M = 4.08$ ). Another significant cost finding was participants' overall negative view of the evaluation process. The mean score when participants were asked to rate the effect of the evaluation process of their effectiveness was 2.92 ( $M = 2.92$ ), indicating a negative to no effect.

**Research Question 4:** *Do special education teachers' perceptions of teacher evaluation systems vary by:*

- a. gender*
- b. job categories, and*
- c. years of experience.*

The inferential analysis conducted to compare participants' perceptions across the gender, job categories, and years of experience revealed the perceptions did not differ across those categories. No statistically significant differences were found for the independent *t*-tests for gender or the ANOVA tests for job categories and years of experience. The findings suggest that regardless of their gender, years of experience, or job categories, participants had similar perceptions about using the existing evaluation systems to rate special education teacher effectiveness.

#### **Understanding of evaluation system and components.**

One of the findings from the study related to the level of understanding special education teachers had about the evaluation system's components and how they are combined to compute summative performance ratings. Regardless of their gender, years of experience, or job description, the findings show that special education teachers did not have a deep understanding of the evaluation system and its components. According to Benedict et al. (2013), it is imperative that special education teachers understand evaluation methods being used to evaluate teacher performance in their districts. Not only will knowledge about how they are assessed help special education teachers in planning impactful lessons for their students, it will also help them perform well in formal evaluations. Benedict et al. (2013) noted that when special education teachers have clarity regarding evaluation expectations in their district, they can adjust their practices to

showcase their strengths and expertise. In addition, special education teachers who possess adequate knowledge about the evaluation system and how it is implemented “will be better equipped with the skills necessary to approach their evaluation with confidence” (Benedict et al., 2013, p. 67).

Studies that solely examined special education teachers’ levels of understanding of evaluation systems and their impact of their practice and professional development needs are virtually non-existent. In their study of teachers’ perceptions about the Chicago Public School REACH evaluation framework involving both general and special education teachers, Jiang et al. (2015) found that new teachers reported having a higher level of clarity of the evaluation system than more experienced teachers, and elementary teachers had higher understanding than high school teachers. Jiang et al. (2015) further found that special education teachers had lower clarity than their general education colleagues. Teachers’ levels of clarity were also found to be moderately related to perceptions about their evaluator, and strongly related to how they perceived the quality of professional development in their schools. In contrast with Jiang et al.’s (2015) finding about the relationship between years of experience and understanding of the evaluation system, this study did not find differences between new and experienced special education teachers relative to their understanding of the evaluation system. Perhaps a bigger sample size would have yielded a different outcome. Admittedly, participants in this study reported not having a higher overall understanding or clarity about the evaluation system in their district, similar to Jiang et al.’s (2015) finding that special education teachers reported significantly lower clarity than non-special educators.



## **Perceptions about Roles, Responsibilities, and Contexts**

Participant perceptions about the practicality of the evaluation system in relation to their roles, responsibilities, and contexts, revealed a number of findings. The findings were disaggregated into areas of evaluator, student growth, and feedback perceptions.

**Perceptions about evaluators.** Irrespective of their gender, years of experience, and job categories, participants were somewhat positive about their evaluators' ability to assess their performance and to do so in a fair and unbiased manner. Although participants were not fully positive, the finding is still significant. It suggests that, overall, participants believed that their evaluators had the needed skills to rate their roles, responsibilities, and contexts if the evaluators were given rubrics designed for special education settings. This is in contrast with Guartico's (2016) finding where the majority of participants perceived their administrators as lacking adequate training to evaluate their practice. Admittedly, participants in the Guartico (2016) study were drawn from the elementary level and from school districts that did not use the Marzano Evaluation Framework. It is not known for sure whether those factors contributed to the contrasting perceptions.

Some of the open-ended responses shed some light on why participants were not fully positive about their evaluators. Some participants believed that their administrators were not knowledgeable about special education classroom practices and teaching methods and therefore could not accurately assess their performance or practice. One participant's (ID #34) comment perhaps encapsulates the reservations expressed:

Special educators are often evaluated by administrators who have no background or education in the dynamics of special education teachers, as such it is almost impossible to get a fair and accurate representation of special educators performance in the class.

Studies about evaluators' ability to rate special education teacher performance have often examined the issue from the viewpoint of administrators. Glowacki (2016) found that elementary school administrators reported being confident in evaluating special education teachers. Significant differences were, however, found between administrators with special education certification and those without. Administrators with special education knowledge reported higher levels of effectiveness with regard to giving relevant feedback to special educators. Steinbrecher et al. (2015) also found that elementary principals lacked knowledge about special education teacher behaviors that operationalized evidence-based practices in special education classroom observations. Even in studies where principals reported feeling more confident in their ability to evaluate special education teachers (e.g., Lawson & Knollman, 2017), they still indicated finding it challenging to evaluate special education teachers in self-contained settings.

**Student growth measures as indicator of effectiveness.** One issue that has consistently generated debate and discussion is the use of student growth measures in special education teacher performance evaluation. Participants were opposed to using student growth metrics as a measure of their performance because they believed that the measures are not a reflection of their students' abilities or learning. Several participants expressed their disagreement with using student growth indicators in their open-ended responses. One participant (ID # 11) in a self-contained setting, whose students take the alternate assessment administered to students with severe disabilities, opposed student growth measures because they do not reflect special education teachers' teaching ability or students' capabilities:

State testing (FSAA) is useless for students with severe cognitive deficits and is not reflection of ESE teachers' ability to teach. Students are often not even able to respond due to the severity of their intellectual disability. It is a waste of time and cruel to make those students sit through hours of individual testing. They do not show improvement on the test and the test does not indicate their needs or abilities or growth. I do not think people realize that the FSAA takes 3-4 hours to administer to EACH student and input

scores. This is what teachers are required to do instead of teaching for almost an entire month.

Another participant (ID # 19) commented that the assessments are problematic because they do not match the skillset students need for better outcomes.

The testing we give to special education students is especially a problem; as an ASD Teacher, I am aware that my students interact with the world and learn much differently than a general education student. My students should not be tested in the same manner, but they are. I am required to read them long passages that they do not understand, and they are required to answer using long statements in which they may only understand one word. It shows a huge lack of awareness and support for these students.

In spite of opposition to student growth measures being used as indicators of effectiveness, participants agreed that they had adjusted their teaching practices with the view to improving the performance of their students on the assessments. The adjustments were possibly made for the sake of complying with requirements of the evaluation rubric or for the purpose of improving their own student growth scores. This possibility can be deduced from one participant who said, "...Even though I think that the assessment process for ESE teachers is flawed I have changed my teaching to meet the observation rubric." There are no studies that have examined how special education teachers have responded or reacted to the inclusion of student growth metrics as part of the evaluation matrix.

Studies have pointed out the negative perceptions of teachers regarding the use of student growth in evaluations (Jiang et al., 2017). In the case of special educators, researchers have consistently pointed out the problematic nature of using student growth measures such as VAM scores to estimate special education teacher effectiveness. Participants' perceptions align with some of the reasons why researchers are opposed to student growth indicators in special education teacher evaluations. The reasons include low scores of students with disabilities, number of students included in VAM computations (Brownell & Jones, 2015; Sledge & Pazey,

2013), mismatch between assessments and abilities of students with special needs, and the diversity of roles assigned to special education teachers (Steinbrecher et al., 2014). While the 2015 passage of ESSA offers states and school districts some flexibility regarding the use of student growth metrics in teacher evaluations, Close et al. (2018) reported that there are still states, like Florida, that encourage the use of VAMs and some school districts continue to use them as in the case of the school district in this study.

### **Responsiveness of evaluation frameworks to roles, responsibilities, and contexts.**

One significant finding from the study was participants' perceptions about the appropriateness of the rubric used to assess their roles, responsibilities, and practices as special educators.

Regardless of their gender, years of experience, or job descriptions, respondents were of the view that the instrument used to evaluate them was inappropriate for special education teaching and practice. One participant (ID #48) was of the view that the rubric is subjective because it does not clearly define behavior expectations for special education teachers or consider their diverse roles and responsibilities.

The evaluation process is..subjective in nature...better defined evaluation processes are needed specifically for special ed teachers since their role is more diversified than the gen ed teacher. Nowhere does the process evaluate the quality or effectiveness of IEP writing, contact with parents and staff, or coordination of various other activities that related to not only job performance but effectiveness as it relates to school operations.

Another participant (ID # 31) was of the view that even the language used gives an indication that the designers of the evaluation rubric simply did not have the roles, responsibilities, contexts of special educators or the needs of their students in mind when they designed the rubric. The participant said,

The new Marzano framework leaves Special Education Educators out completely. "Close the achievement gap" is one of our areas. Seriously?? Which achievement gap am I trying to close-the fact that my students are 21 years old physically yet 2 years old

mentally? The gap between the fact that they are at an ESE Center as opposed to a Gen. Ed site? It's as if the Marzano people forgot about special education all together.

The views expressed by participants in the study were similar to what other studies have revealed about how special education teachers view the existing evaluation frameworks adopted by states and school districts. In a study that examined elementary special education teachers' perceptions about the design and implementation of the evaluation process across three school districts, Guartico (2016) reported that special education teachers viewed certain components of the evaluation rubrics as inapplicable to their roles and responsibilities. Mrla's (2016) study of special educating teachers' perceptions about the relevance of the State of Arkansas' Teacher Excellence Support System (TESS) teacher evaluation rubric revealed reported similar findings. On the other hand, Lawson (2015) found that special education teachers were divided regarding the appropriateness of RESET evaluation tool, an adapted version of the Danielson Framework for special education teachers. While some perceived the tool as appropriate, others perceived it as inappropriate citing the failure of the instrument to consider special education teachers' non-instructional roles such as parent communication, IEP-related responsibilities, and service coordination in the evaluation process. Several studies (e.g. Jones & Brownell, 2014; CEC, 2012; Crowe et al., 2017; Johnson & Semmelroth, 2014; Jones, 2016; Sledge & Pazey, 2013) have also examined the applicability of existing evaluation tool to special education teacher roles and practice and concluded that there are challenges associated with using existing tools to evaluate special education teachers.

**Different evaluation system for special education teachers.** Participants in the study overwhelmingly recommended a separate evaluation tool for special education teachers. The analysis of data showed that participants were not sure of the relevance of the feedback they

received from the evaluation process. The findings also indicate that participants were not sure about using the feedback received to improve their practice and teaching roles and responsibilities. The finding regarding a recommendation for separate evaluation tool aligns with Guartico's (2016) study, which found that special education teachers from three different districts, and evaluated with different tools, called for different evaluation measures that would be sensitive to special education classrooms and settings. While researchers agree on the challenges associated with using instruments not validated for special education teachers to evaluate them, there is yet to be consensus on how the challenges should be addressed. One group of researchers (e.g. Brown & Jones, 2015; Holdheide et al., 2010; Johnson & Semmelroth, 2014) would not subscribe to the idea of separate evaluation systems. Instead, this group of researchers recommend modifications and adaptations of existing rubrics to include competencies and behavior indications of effective special education teaching and practice. The other group of researchers (e.g., Barnes et al., 2018; Crowe et al., 2017; Sledge & Pazey, 2013) support the call for differentiated evaluation measures. The findings this study and Guartico (2016) give some indication that special education teachers at all levels (elementary to high school) and diverse settings support the idea that it may be time to consider developing separate, validated evaluation systems that are responsive to the needs of special education settings and contexts.

**Evaluation feedback did not inform professional development choices.** Another important finding from the study related to the extent to which participants reflected on the feedback received and used it to inform their professional development planning and choices. Overall, the findings revealed that the evaluation process did not encourage participants to actively reflect on their teaching practice or make changes as a result. The findings further indicate that the evaluation process had limited influence on participants' professional

development choices. The fact that participants did not feel inclined to reflect on their teaching or make changes to the pedagogy and practice perhaps explains why the process did not inform their professional development planning and choices. Participants did not view the process and professional development activities associated with it as practical to their practice. One participant (ID #38) expressed it succinctly, “I...strongly believe that the majority of staff development or training organized by the county or school do not pertain to my classroom and I cannot use most of the concepts introduced.” Another participant (ID #52) explained, “...the working relationship with my team and leadership have had an impact on my practice and professional development choices than the evaluation/observation rubric.” This participant believed collaboration with colleagues and leadership would better meet their practice needs than professional development tied to the evaluation process.

This finding is consistent with Guartico’s (2016) study in which elementary self-contained special education teachers indicated that the evaluation process did not influence their professional development choices. Even when paired with monetary compensation and leadership opportunities, Guartico (2016) reported that special education teachers in the study were not motivated to take advantage of professional development initiatives tied to the evaluation process. They would rather collaborate with colleagues to learn and share ideas that could improve their practice. If participants did not consider the feedback received as relevant when making professional development choices, it is imperative to find out how they developed their deliberate practice plans. While there is no evidence to suggest that participants did not consult or reflect on the evaluation feedback when writing their deliberate practice plans, it is possible that participants viewed the process as a mere formality that they had to comply with.

According to Brownell and Jones (2015), one of the functions of effective teacher evaluation systems is to guide school districts in how to direct their professional development initiatives that help special education teachers identify the requisite skills and knowledge to improve their classroom practice and instruction. The CEC (2013) position paper on special education teacher evaluation indicated, however, that in most cases, teacher evaluation systems are not related to professional development needs of teachers. Researcher have stressed that any valid evaluation system must provide useful feedback that fosters improvement in professional development and practice for special education teachers (e.g., Johnson & Semmelroth, 2014; Semmelroth, Johnson, & Allred, 2013). Woolf (2018) reported that teacher educators and school administrators emphasized ongoing professional development and reflection as critical to the function of a special education teacher.

**Impact of evaluations on special education teachers.** The findings from the study further revealed participants' perceptions about the cost or impact of the evaluation system of their well-being. Participants were of the view that the evaluation process is not worth all the effort it requires mainly due to the level of stress and anxiety associated with it. Participants further revealed that the evaluation process had not been impactful on their effectiveness as special education teachers. Jiang et al. (2015) found that teachers reported increased stress and anxiety levels due to teachers' perceptions of the associated costs by way of apprehension, time, and energy expended on the process, although the feelings decreased over time. For special education teachers, there are no studies that have investigated the how they have responded to being evaluated with systems designed for general education teachers or how they have performed relative to their practices (Holheide, 2015). The findings in this study provide some vital insights into the impact of the evaluation process on special education teachers.



## **Implications for Practice and Policy**

Teacher evaluation systems have taken the center stage in the efforts to ensure that all students are taught by effective teachers. To help distinguish effective teachers from ineffective ones, teacher evaluation systems should clearly identify what constitutes effectiveness in diverse classroom settings and provide relevant feedback that delineates the knowledge and skills teachers need to improve their performance and practice. For special education teachers, Brownell and Jones (2015) identified at least three functions teacher evaluation must serve: (1) provide information about distribution of special education teachers in states and districts, (2) help special education teachers identify skills and knowledge needed to improve their pedagogy and practice and professional development planning, and (3) identify effective special education teachers for recognition and ineffective ones who do not qualify for special education careers. Achieving these objectives with existing evaluation systems, however, has been fraught with challenges emanating from the complex nature of special education teacher roles, nature of special education teaching and classrooms contexts, absence of validated evaluation rubrics specific to special education practice, and administrators' lack of knowledge of special education practices and settings. Added to these challenges is the lack of empirical research documenting how special education teachers perceive these systems and their impact on their effectiveness and professional development needs.

The findings from the study suggest that participants did have a deep understanding of the evaluation process, especially how the different components contribute to the overall performance ratings. The findings show participants had slightly favorable views about their evaluators ability to evaluate them and to do so fairly. On the other hand, the findings reveal that participants did not perceive the evaluation system as practical in the areas of student

performance metrics, feedback received, responsiveness to job roles and responsibilities, and professional development choices. A significant finding was the strong call for separate evaluation system for special education teachers. Another important finding was the perception about the emotional cost of the evaluation process. These findings have implications for special education practice and policy at the district, state, and federal levels.

**Implications for school districts.** The success or failure of policies, such as the adoption of an evaluation framework by states or districts, depends, to a large extent, on the perceptions and reactions of final implementation agents (Doyle and Ponder, 1977; Fullan, 2001; Muncey & McQuillan, 1996). School districts have adopted evaluation systems to help distinguish between effective and ineffective teachers and to empower “teachers with the tools and resources necessary to grow their practice” (Carbaugh, Marzano, & Toth, 2017, p. 3). It is imperative, therefore, that teachers upon whom evaluation measures are being implemented, have a clear understanding of the evaluation process, expectations, criteria for performance, and the components used to judge their effectiveness. If special education teachers do not have clarity about the criteria for performance determination, then, questions remain about how they plan and teach to meet the requirements of the evaluation system. Apart from being the gatekeepers of evaluation policies, meeting the expectations or otherwise has implications for special education teacher growth, remuneration, and ultimate retention in the profession.

Districts can organize districtwide professional development sessions to help teachers understand the components and expectations of adopted evaluation tools. The professional development session could be tailored to the needs of subgroups of teachers such as special education teachers. The findings from the study suggest that districts may need to embed more school and team level opportunities, such as lunch and learns and professional learning

communities, that allow for small group and personal interactions to reinforce teacher knowledge and understanding of the evaluation process in the district. Districts can also embed an evaluation component into their mentoring programs as part of the onboarding process to allow experienced administrators and mentors to support new special education teachers' understanding of the evaluation system and accompanying processes (Israel et al., 2014).

Another implication from the findings of the study is the need for district to adopt evaluation systems that addresses the need of all groups of teachers. The findings clearly indicate that participants did not perceive the existing evaluation framework as practical for their needs. Whether districts adopt commercially available systems such as Danielson's FFT or the Marzano Framework, or design their own system, it is imperative for districts to consider that needs special education teachers given the uniqueness of the roles, responsibilities, and contexts. Granted, districts that may be constrained by contractual obligations with developers of adopted systems need to work with developers to adapt the tool to suit special education settings as has been done by some states and districts (Brownell & Jones, 2015; Holdheide, 2015). Regardless, special education teacher responsibilities such as IEP-related responsibilities, service coordination, and collaboration with teachers and families need to be incorporated and evaluated. Districts that choose to design their own systems need to align their rubrics with evidenced-based practices and professional standards revealed in special education research and those espoused by special education professional organizations such as the CEC.

The findings further highlight the need for districts to adopt evaluation systems that are able to provide relevant feedback to special education teachers for the purposes professional development planning. Districts need to invest in research that solicits the views of teachers regarding the nature and relevance of feedback they receive from their annual evaluations.

Districts can use the feedback to develop target professional development options that address the needs of various groups, such as special education teachers. The need to ensure better academic outcomes for students with disabilities necessitates that special education teachers receive useful feedback and subsequent professional development opportunities to hone their skills and practice.

For special education teachers, the findings from the study suggest a need to advocate for their needs in the evaluation process. Participants in the study perceived their evaluators as capable of assessing their effectiveness in a fair manner contrary to the findings of other studies that revealed special education teachers' negative views towards their evaluators' abilities and knowledge of special education practice. It suggests that participants had good relationships with their evaluators and trusted them to have their interests at heart. In view of the fact that special education teacher voices have largely been absent in the conversation about evaluation process and its impact on their practice, special education teachers capitalize on their relationships with administrators to start a conversation about the inclusion of the voices and inputs starting with their evaluators at each school site.

**Implications for state education agencies.** The 2015 passage of ESSA has reduced federal oversight and instead increased state control in terms of the design and implementation of accountability measures such as teacher evaluation systems. According to Crowe et al. (2018), states have abandoned the one-size-fits all approach to teacher evaluations. Instead, many state ESSA plans reveal guidelines to help states choose and implement evaluation models. There is, however, wide variation in the firmness of the guidelines (Crowe et al., 2018). In Florida, for example, the state has given a set of core standards known as the Florida Education Accomplished Practices (FEAPs) to guide public schools and teacher education programs on

knowledge and skills needed by effective educators in the state (Florida Department of Education, 2019). School districts in Florida, including one where participants were drawn from, have made efforts to align their evaluation systems, including adopted ones such as the Marzano Focused Model, to the FEAPs standards. The findings from the study suggest that the needs of all teachers such as special education are not being met. States need to provide uniform guidance provided addresses the unique roles and responsibilities of special education teachers to school districts. States can partner with researchers to design or adapt evaluation systems that incorporate evidence-based special education practices that can improve outcomes for students with disabilities. The partnership between researchers from Boise State University and the State of Idaho Department of Education to develop the RESET tool (Johnson & Semmelroth, 2014) is a worthy example.

**Implication for federal education policy.** Although ESSA has reduced the level of federal involvement and influence in the design and implementation of teacher evaluation systems, the federal role in ensuring that students with disabilities have access to high quality and effective special education teachers remains largely intact and critical. The US Department of Education (USDOE), through its Office of Special Education Programs, has for years committed to improving access and better outcomes for students with disabilities by providing financial support for leadership and personnel development to states and school districts. One of OSEP's flagship programs is the personnel preparation program that awards funding to institutions of higher education to train certified leadership and instructional personnel to support children with disabilities. One of the purposes the personnel preparation grants is to ensure that personnel trained through the program have the needed skills and knowledge to serve children with disabilities (USDOE, 2016). To further ensure the effectiveness of special

education teachers, the USDOE should provide funding for research into designing and implementing evaluation systems that fairly and equitably address the responsibilities, roles, and competencies of special education teachers. The need to fund research into this avenue of research is pertinent given the persistent problem of attrition in special education.

### **Recommendations for Future Research**

The findings of this study contribute to the emerging research on special education teacher evaluation, especially in terms of how special educators perceive the existing evaluation frameworks relative to their roles, responsibilities, and professional development needs. Ongoing research on the evolving role of teacher evaluation in determining special education teacher effectiveness and ensuring equitable education outcomes is definitely warranted. The following recommendations are offered for future research:

1. The current study involved secondary level special education teachers drawn from one school district. At least one other study (Guartico, 2016) has explored the perceptions of elementary level special educators. Future studies involving participants from K-12 levels and from multiple school districts with different evaluation systems and tools is needed provide a comprehensive understanding of the perceptions about existing evaluation systems. Another group that has not been studied is special education teachers in virtual settings. Studies about their perceptions relative to how their effectiveness is evaluated merits attention.
2. The study examined their perceptions through quantitative approaches. Future studies can employ mixed method design that can solicit both quantitative and qualitative data for in-depth understanding and perspectives about special education teacher perceptions of the evaluation systems.

3. Participants in the study overwhelmingly recommended the adoption of separate evaluation system for special education teachers as a result of their perception that existing framework is not differentiated for their practice. Further research should examine the validity of approaches and instruments that are being designed by some researchers for special education classrooms (e.g., Barnes, et al., 2018; Johnson & Semmelroth, 2015) and how special education teacher perceive such instruments.
4. Research examining how states and districts have revamped their evaluation systems to assess the competencies of special education teachers in the ESSA era, including the extent and in what capacities student growth measures are being used, is also worthy of attention.
5. Studies are also needed to examine the emotional and psychological impact of teacher evaluation systems on special education teachers.
6. Researchers who would like to use the instrument adopted for the study may need to revise or modify it to align with their research questions and purpose.

### **Limitations of Study**

The following are some limitations of the study:

- This study was conducted using a survey that attained a response rate of 26.8%. As in any quantitative study, the sample size has implication for the representativeness of findings.
- The findings of the study may not be generalizable to other special education teachers in districts that do not evaluate their teachers with the same rubric as the district in which the study will be conducted. The district from which participants were drawn evaluates its

teachers using the Marzano Focused Teacher Evaluation Model. Therefore, findings from the study would have to be interpreted within that context.

- The choice of methodology impacted the ability to obtain in-depth insights into how special education teachers perceive evaluations systems, including how contextual factors at each school site relate to perceptions. Data were collected through census survey and did not include multiple sources of data such interviews, documents, and school-related information that would have provided contextual and background information to interpret the findings.
- The timing of the study may have impacted the response rate (26.8%), although it is not known for sure. The timeframe for data collection, April 1 – 26, 2019 coincided with the ‘testing season’ when schools in Florida administer the state’s standardized assessments. Teachers who proctor the tests typically have limited access to their work emails due to restrictions imposed on proctors. Admittedly, participants could complete the survey at any time and place including after-school hours.

### **Summary of Chapter Five**

This chapter discussed the findings of the study within the context of the framework that guided that study and the extant literature on the subject of special education teacher evaluation. Implication of the findings for the district, state, and federal policies were also discussed. For districts, the findings show the need to: (1) ensure their special education teacher understand adopted evaluation systems through districtwide and school level professional development opportunities, as well on-boarding mentoring efforts, and (2) adopt or design evaluation systems that address the complex and varied responsibilities of special education teachers. Implication for state education agencies is to provide a uniform framework to guide districts in their adoption of



systems that will address the practice and needs of special educators. For federal policymakers, it is vital that the USDOE and OSEP provide funding for research into the issue of special education teacher evaluation. Recommendations for future research include studying perceptions of special education teachers across all grade levels and from multiple school districts with different evaluation systems. Future research can also employ mixed-method approaches, evaluate the validity of emerging tools designed for special education contexts, and assess the emotional toll of existing systems on special education teachers.

### **Conclusion**

The challenges associated with evaluating special education teachers with existing evaluation systems are well-known (Brownell & Jones, 2015; Holdheide, 2015; Sledge & Pazey, 2013). Researchers continue to debate the best approaches to measure special education teacher effectiveness, an issue at the core of the teacher evaluation efforts. All through the debates, special education teachers have been omitted from large scale studies examining the impact of teacher evaluation reforms (Jones & Brownell, 2014; Sledge & Pazey, 2013) and their perceptions, including how they have fared, have not been studied. Three significant problems necessitated this study: (1) the challenges surrounding evaluation of special education teacher effectiveness, (2) the paucity of research documenting special education teacher views, perceptions, voices, and experiences in terms of how they have fared under existing evaluations systems, and (3) the persistent problem of special education teacher attrition within the context of high-stake teacher evaluations. Findings from the study underscore the importance of including special education teacher voices in the discussions surrounding the best approaches to evaluate the complex roles, responsibilities, and practices special educators. The findings stress the need for researchers and policy makers to give considerable attention to the issue of special

education teacher evaluation as it has implications for student outcomes as well as the persistent problem of attrition that has bedeviled the field for several years.

## REFERENCES

- Ary, D., Jacobs, L.C., & Sorensen, C. (2010). *Introduction to research in education*. Belmont, CA: Wadsworth, Cengage Learning.
- Bacon, J. (2015). The impact of standard-based reform on social education and the creation of the 'dividual'. *Critical Studies in Education*. 52(3), 366-383. Doi: 10.1080/175508487.2105.979845
- Barnes, T. N., Cipriano, C., Flynn, L., Rivers, S. E., & Xu, W. (2018). Validating the recognizing excellence in learning and teaching (RELATE) tool for special education classrooms. *The Journal of Experimental Education*, , 1-15.  
doi:10.1080/00220973.2018.1465383
- Benedict, A. E., Thomas, R. A., Kimerling, J., & Leko, C. (2013). *Trends in teacher evaluation: what every special education teacher should know*. United States: COUNCIL FOR EXCEPTIONAL CHILDREN. Retrieved from <http://ezproxy.lib.usf.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edsbl&AN=RN332213799&site=eds-live>
- Boe, E.E. (2014). Teacher demand, supply, and shortage in special education: A national perspective. *Handbook of research on special education teacher preparation*. Edited by Paul T. Sindelar, Erica D. McCray, Mary T. Sindelar, Benjamin Lignugaris/Kraft. Routledge: New York.

- Boote, D. N., & Beile, P. (2005). Scholars before researchers: On the centrality of the dissertation literature review in research preparation. *Educational Researcher*, 34(6), 3 – 15.
- Brownell, M.T. & Jones, N.D. (2015). Teacher education in special education: Approaches, supporting research, and challenges confronted. *Journal of Special Education Leadership*, 28(2), 63 – 73.
- Carbaugh, C., Marzano, R., & Toth, M. (2017). The Marzano focused teacher evaluation model: a focused, scientific-behavioral evaluation model for standards-based classrooms. *Learning Sciences Marzano Center*. Retrieved on 5/13/2019 from [www.learningsciences.com](http://www.learningsciences.com)
- Check, J., & Schutt, R. K. (2011). *Research methods in education*. Los Angeles, CA: Sage.
- Close, K., Amrein-Beardsley, A., & Collins, C. (2018). *State-Level Assessments and Teacher Evaluation Systems after the Passage of the Every Student Succeeds Act: Some Steps in the Right Direction*. Boulder, CO: National Education Policy Center. Retrieved on 7/10/18 from <http://nepc.colorado.edu/publication/state-assessment>.
- Cohen, J., & Goldhaber, D. (2016). Building a more complete understanding of teacher evaluation using classroom observations. *Educational Researcher*, 45(6), 378–387.
- Council for Exceptional Children (2012). *The Council for Exceptional Children's position on special education teacher evaluations*. The Council for Exceptional Children.
- Creswell, J.W. (2003). *Research design: Qualitative, quantitative and mixed method approaches*. Thousand Oaks, CA: Sage.
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*. Los Angeles, CA: Sage.

- Creswell, J.W. (2015). *Educational research: planning, conducting, and evaluating qualitative and quantitative research, 5<sup>th</sup> Edition*. Upper Saddle River: NJ, Pearson.
- Crowe, C. C., Rivers, S. E., & Bertoli, M. C. (2017). Mind the gap: Accountability, observation and special education. *Assessment in Education: Principles, Policy & Practice*, 24(1), 21-43. doi:10.1080/0969594X.2015.1114913
- Danielson, C., & McGreal, T. (2000). *Teacher Evaluation to Enhance Professional Practice*. Alexandria, VA: Association for Supervision & Curriculum Development.
- Darling-Hammond, L. (2013). *Getting teacher evaluation right: What really matters for effectiveness and improvement*. New York: Teachers College Press.
- Doabler, C.T., Nelson, N.J., Kosty, D.B., Fien, H., Baker, S.K., Smolkowski, K., & Clarke, B. (2014). Examining teachers' use of evidence-based practices during core mathematics instruction. *Assessment for Effective Instruction*, XX(X) 1 – 13
- Doerr, S.E. (2012). *Charlotte Danielson's theory of teacher evaluations: a quantitative study of teachers' perceptions on the four domains*. Dissertation.
- Doyle, W., & Ponder, G. (1977). The practicality ethic in teacher decision-making. *Interchange*, 8(3), 1–12.
- Editorial Projects in Education Research Center. (2015, September 3). Issues A-Z: Teacher Evaluation: An Issue Overview. *Education Week*. Retrieved 02/12/2016, from <http://www.edweek.org/ew/section/multimedia/teacher-performance-evaluation-issue-overview.html/>

Elliot, S.N., Roach, A.T. & Kurz, A (2014). Evaluating and advancing the effective teaching of special educators with a dynamic instructional practices portfolio. *Assessment for Effective Intervention*, XX(X) 1 – 16

Florida Department of Education (2018). *Identification of critical shortage areas*. Retrieved online at <http://www.fldoe.org/core/fileparse.php/7584/urlt/CTSA1819.pdf>

Florida Department of Education (2019). *The Florida Accomplished Practices (FEAPs)*. Retrieved online from <http://www.fldoe.org/teaching/professional-dev/the-fl-educator-accomplished-practices.stml> on May 26, 2019.

Fullan, M. (2001). *The new meaning of educational change* (3rd ed.). New York: Teachers College, Columbia University.

Glowacki, H.L. (2013). *Illinois Elementary Principals' perceptions of the teacher evaluation process of special education teachers*. Dissertation.

Glowacki, H., & Hackmann, D. G. (2016). The effectiveness of special education teacher evaluation processes: Perspectives from elementary principals. *Planning & Changing*, 47(3), 191-209. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=eue&AN=123507954&site=ehost-live>

Gordon, R., Kane, T., & Staiger, D. (2006). Identifying effective teachers using performance on the job (Discussion Paper 2006–01). Washington, DC: Brookings Institution.

- Grissom, J. A., & Youngs, P. (2016). Making the most of multiple measures. *Improving teacher evaluation systems: making the most of multiple measures*. Edited by Jason A. Grissom & Peter Youngs. Teachers College Press: New York.
- Groves, R.M., Fowler, F.J. Couper, M.P., Lepkowski, J.M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology*, 2<sup>nd</sup> Edition. Hoboken, N.J.: John Wiley & Sons.
- Guarino, C., Santibanez, L., & Daley, G. (2006). Teacher Recruitment and Retention: A Review of the Recent Empirical Literature. *Review of Educational Research*, 76(2), 173–208.
- Guartico, R. A. (2017). *Special education teachers' perceptions of the design and implementation of the evaluation process*. (2016-47715-230). Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2016-47715-230&site=ehost-live>
- Hazi, H.M. & Rucinski, D.A. (2009). Teacher evaluation as a policy target for improved student learning: a fifty-state review of statute and regulatory action since NCLB. *Education Policy Analysis Archives* 17(5), 1 – 22
- Holdheide, L. (2015). Same Deabte, New Opportunities: designing teacher evaluation systems that promote and support educators in practices tha advance all students' learning. *Journal of Special Education Leadership* 28(2), 74 – 81
- Holdheide, L.R., Goe, L., Croft, A., & Reschly, D.J. (2010). Challenges in Evaluating Special Education Teachers and English Language Learner Specialists. *National Comprehensive Center for Teacher Quality*.

Israel, M., Kamman, M. L., McCray, E. D., & Sindelar, P. T. (2014). Mentoring in action: the interplay among professional assistance, emotional support, and evaluation. *Exceptional Children*, 81(1), 45-63.

Jiang, J. Y., Spote, S. E., & Luppescu, S. (2015). Teacher perspectives on evaluation reform: Chicagos REACH students. *Educational Researcher*, 44(2), 105.

Johnson, E.S., Crawford, A., Moylan, L.A., & Ford, J.W. (2016). Issues in evaluating special education teachers: challenges and current perspectives. *Texas Education Review*, 4(1), 71 – 83.

Johnson, E., & Semmelroth, C. L. (2014). Special education teacher evaluation: Why it matters, what makes it challenging, and how to address these challenges. *Assessment for Effective Intervention*, 39(2), 71-82. doi:10.1177/1534508413513315

Johnson, E., & Semmelroth, C. L. (2014a). Introduction to AEIs special issue on special education teacher evaluations. *Assessment for Effective Intervention*, 39(2), 67-70. doi:10.1177/1534508413511489

Jones, N.D. (2016). Special education teacher evaluation: An examination of critical issues and recommendations for practice. *Improving teacher evaluation systems: making the most of multiple measures*. Edited by Jason A. Grisson & Peter Youngs. Teachers College Press: New York.



Jones, N. D., & Brownell, M. T. (2014). Examining the use of classroom observations in the evaluation of special education teachers. *Assessment for Effective Intervention*, 39(2), 112-124. doi:10.1177/1534508413514103

Joyce, J., Harrison, J.R., & Murphy, D. (2016). Evaluating students with disabilities and their teachers: use of student learning objectives. *Teacher College Record* 118(140311), 1- 22.

Keenan, M. (2018). *Perception*. Salem Press Encyclopedia of Health. Retrieved from <http://search.ebscohost.com.ezproxy.lib.usf.edu/login.aspx?direct=true&db=ers&AN=109057209&site=eds-live>

Kelchtermans, G. (2017). ‘Should I stay or should I go?’: unpacking teacher attrition/retention as an education al issue. *Teachers and Teaching*, 23:8, 961 – 977, DOI: 10.1080/13540602.2017.1379793

Lawson, J (2014). Value-added modeling: challenges for measuring special education teacher quality. *UCLA Journal of Education and Information Studies*, 10(1)

Lawson, J. E. (2015). *Evaluating special education instructional practices using observation rubrics: Investigating the reliability of school administrator ratings*. (2016-26522-154). Retrieved from <http://ezproxy.lib.usf.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2016-26522-154&site=eds-live>

Lawson, J. E. & Knollman, G. A. (2017). Evaluating special education teachers: School administrators' perceptions of the process. *Journal of Special Education Leadership*, 30(1), 6-18. Retrieved

from <http://search.ebscohost.com/login.aspx?direct=true&db=eue&AN=122350921&site=ehost-live>

- Milanowski, A. (2017). Lower performance evaluation practice ratings for teachers of disadvantaged students: Bias or reflection of reality? *AERA Open*, 3(1), 2332858416685550. doi:10.1177/2332858416685550
- Moran, R.M.R. (2017) The Impact of a High Stakes Teacher Evaluation System: Educator Perspectives on Accountability, *Educational Studies*, 53:2, 178-193, DOI: 10.1080/00131946.2017.1283319
- Muncey, D., & McQuillan, P. (1996). *Reform and resistance in schools and classrooms*. New Haven, CT: Yale University Press.
- National Council for the Accreditation of Teacher Education (2010). *Transforming Teacher Education Through Clinical Practice: A National Strategy to Prepare Effective Teachers*. Report of the Blue Ribbon Panel on Clinical Preparation and Partnerships for Improved Student Learning.
- Pizmony-Levy, O., & Woolsey, A. (2017). Politics of education and teachers' support for high-stakes teacher accountability policies. *Education Policy Analysis Archives*, 25(87). <http://dx.doi.org/10.14507/epaa.25.2892>

- Quigney, T. A. (2010). ESEA's proposed pay-for-performance option: Potential issues regarding the evaluation of special educators. *Academic Leadership*, 8(4), 1-8. Retrieved from [http:// contentcat.fhsu.edu/cdm/compoundobject/col lection/p15732coll4/id/537](http://contentcat.fhsu.edu/cdm/compoundobject/col lection/p15732coll4/id/537)
- Sledge, A., & Pazez, B. L. (2013). Measuring teacher effectiveness through meaningful evaluation: Can reform models apply to general education and special education teachers? *Teacher Education & Special Education*, 36(3), 231.
- Steinbrecher, T. D., Selig, J. P., Cosbey, J., & Thorstensen, B. I. (2014). Evaluating special educator effectiveness: Addressing issues inherent to value-added modeling. *Exceptional Children*, 80(3), 323-336. doi:10.1177/0014402914522425
- Sutcher, L., Darling-Hammond, L., & Carver-Thomas, D. (2016). *A coming crisis in teaching? Teacher supply, demand, and shortages in the U.S.* Palo Alto, CA: Learning Policy Institute
- Rasul, N. (2018). *Elementary Principals' Perceptions of Special Education Teachers' Performance Evaluation Process*. ProQuest Dissertation.
- Tandy, J. A., Whiford, D.K., & Hirth, M.A.(2016). A review of special education teacher (SET) evaluation practices. *Journal of Ethical Educational Leadership*, 3(5), 1 – 16. Retrieved from: <http://www.cojeel.org>
- United States Department of Education (2018). *Teacher Incentive Fund*. Available online at <https://innovation.ed.gov/what-we-do/teacher-quality/teacher-incentive-fund/>

United States Department of Education (2016). *Special education—personnel development to improve services and results with disabilities*. Available online at

<https://www2.ed.gov/programs/osepprep/index.html>. Retrieved on May 26, 2019.

Vannest, K. J., & Hagan-Burke, S. (2009). Teacher time use in special education. *Remedial and Special Education, 31*, 126– 142. doi:10.1177/0741932508327459

Watts, G.W. (2016). An overview of the issues in evaluating special educators: Current challenges and recent developments. *Texas Education Review, 4*(1), 46 – 53.

Widener, R.E. (2011). Evaluating special education teachers: Do we get the job done? A Regional Perspective. *East Tennessee State University*. Retrieved from <http://dc.etsu.edu/cgi/viewcontent.cgi?article=2478&context=etd>

Woolf, S. B. (2015). Special education professional standards. *Teacher Education & Special Education, 38*(4), 276-290. doi:10.1177/0888406414557284

Woolf, S.B. (2018). Critical skills for special educator effectiveness: Which one matter most, and to whom? *Teacher Education and Special Education 1- 15* DOI: 10.1177/0888406418776714

## **APPENDICES**

## **Appendix 1: Survey Instrument**

### **Q16 Informed Consent to Participate in Research Information to Consider Before Taking Part in this Research Study**

Pro # 00039485

Researchers at the University of South Florida (USF) study many topics. To do this, we need the help of people who agree to take part in a research study. This form tells you about this research study. We are asking you to take part in a research study that is called: The Teacher Evaluation Conundrum: Examining the Perceptions of Special Education Teachers. The person who is in charge of this research study is Gordon Brobbey. This person is called the Principal Investigator.

#### **Purpose of the Study**

The purpose of this study is to explore the perceptions of special education teachers regarding the evaluation system used by your district to assess their performance and effectiveness. The study will also allow the researcher to examine the extent which feedback from teacher evaluation systems address special education teacher practice and professional development needs.

#### **Why are you being asked to take part?**

You are being asked to participate in this research study because you have been identified by Pinellas County Schools as a special education teacher whose primary responsibility is to work with students identified as having disabilities. Your views would provide the appropriate insight into how special education teachers perceive the teacher evaluation system used in your district.

#### **Study Procedures**

If you take part in this study, you will be asked to complete an anonymous online survey through Qualtrics through which data will be collected for the purpose of this research. Your completion of the online survey will not in any way be linked to your identity and no personally-identifying information will be collected.

#### **Alternatives / Voluntary Participation / Withdrawal**

You have the alternative to choose not to participate in this research study. You should only take part in this study if you want to volunteer; you are free to participate in this research or withdraw at any time. There will be no penalty or loss of benefits you are entitled to receive if you stop taking part in this study.

#### **Benefits and Risks**

We are unsure if you will receive any benefits by taking part in this research study. This research is considered to be minimal risk.

#### **Compensation**

We will not pay you for the time you volunteer while being in this study.

#### **Privacy and Confidentiality**

We must keep your study records as confidential as possible. It is possible, although unlikely, that unauthorized individuals could gain access to your responses because you are responding online.

Certain people may need to see your study records. By law, anyone who looks at your records must keep them completely confidential. The only people who will be allowed to see these records are: the Principal Investigator, the advising professor, and the University of South Florida Institutional Review Board (IRB).

It is possible, although unlikely, that unauthorized individuals could gain access to your responses. Confidentiality will be maintained to the degree permitted by the technology used. No guarantees can be made regarding the interception of data sent via the Internet. However, your participation in this online survey involves risks similar to a person's everyday use of the Internet. If you complete and submit an anonymous survey and later request your data be withdrawn, this may or may not be possible as the researcher may be unable to extract anonymous data from the database.

#### Contact Information

If you have any questions about your rights as a research participant, please contact the USF IRB at (813) 974-5638 or contact by email at [RSCH-IRB@usf.edu](mailto:RSCH-IRB@usf.edu). If you have questions regarding the research, please contact the Principal Investigator at 813-770-6961 or [gbrobbey@mail.usf.edu](mailto:gbrobbey@mail.usf.edu).

We may publish what we learn from this study. If we do, we will not let anyone know your name. We will not publish anything else that would let people know who you are. You can print a copy of this consent form for your records.

I freely give my consent to take part in this study. I understand that by proceeding with this survey I am agreeing to take part in research and I am 18 years of age or older.

Q1 What is your gender?

- Female
  - Male
  - Prefer not to respond
- 

Q2 Years of experience as a special education teacher?

- 0 - 4 years
  - 5 -9 years
  - 10 years and above
- 

Q3 What is your job description? (Check one)

- Support Facilitator (e.g. VE )
  - Self-Contained ASD
  - Self-Contained InD
  - Self-Contained EBD
  - On-the-Job (OJT) Teachers
  - Resource Room Teacher
  - Other (Please specify) \_\_\_\_\_
-



Q4 How would you rate your understanding of the following parts of your teacher evaluation system?

	Poor (1)	Fair (2)	Good (3)	Very Good (4)	Excellent (5)
The observation rubric or framework used to rate special education teacher professional practice.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
How different assessments are combined to create student growth measure.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
How different observations and student growth measures are combined to determine a summative performance evaluation score.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



Q5 To what extent do the following statements describe your evaluator for this year's evaluation?  
 My evaluator...

	Not At All (1)	A Little (2)	To Some Extent (3)	To A Great Extent (4)
Is able to accurately assess my instruction, roles, responsibilities, and practice.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Understands my strengths and weaknesses as a special education teacher.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Is fair and unbiased.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Understands what is going on in my classroom.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



Q6 To what extent do you agree or disagree with the following statements?

	Strongly Disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly Agree (5)
The rating scales used to evaluate my roles, responsibilities, and practices as a special education teacher are appropriate.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I am given useful feedback by the evaluator	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The feedback I received this year identified specific areas of my instruction, roles, responsibilities, and practice that could be improved.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I have used the feedback I received so far to improve my teaching, roles, responsibilities, and practice.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The feedback I received this year included guidance on how to make improvements to my instruction and special education practice.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I would recommend a different evaluation rubric for special education teachers compared to general education teachers.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q7 The following questions are about measures of student growth based on Florida Standards Assessment (FSA), Florida Standards Alternate Assessment (FSSA), and district-mandated End of Course (EOC) assessments. To what extent do you agree or disagree with the following statements?

	Strongly disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly Agree (5)
The measures of student growth are a fair representation of my students' learning.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The information I got from those assessments will inform my professional development choices.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I have made changes in my teaching and practice in order to improve my students' scores on these assessments.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



Q8 The following questions are about professional development and practice. To what extent do you agree or disagree with the following statements?

	Strongly Disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly Agree (5)
The observation process encouraged me to reflect on my teaching practice.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
My observation ratings will guide my future professional development choices.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
My observation results will strongly influence my future professional development activities.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I have made changes in my teaching and practice as a result of the observation process.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



Q9 To what extent do you agree or disagree with the following statements?

	Strongly Disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly Agree (5)
The evaluation process has increased my level of stress and anxiety.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Overall, the evaluation system takes more effort than the results are worth.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q10 In general, what kind of effect do you think the existing evaluation has had on your effectiveness as a special education teacher?

- Strongly Negative Effect (1)
- Negative Effect (2)
- No effect (3)
- Positive Effect (4)
- Strongly Positive Effect (5)

Q11 Any other comments regarding the evaluation process, the feedback you receive, and the impact on your practice and professional development choices.

---

## Appendix 2: Informed Consent



## **Informed Consent to Participate in Research**

Information to Consider Before Taking Part in this Research Study

**Pro # 00039485**

Researchers at the University of South Florida (USF) study many topics. To do this, we need the help of people who agree to take part in a research study. This form tells you about this research study. We are asking you to take part in a research study that is called: *The Teacher Evaluation Conundrum: Examining the Perceptions of Special Education Teachers*. The person who is in charge of this research study is **Gordon Brobbey**. This person is called the Principal Investigator.

### **Purpose of the Study**

The purpose of this study is to explore the perceptions of special education teachers regarding the evaluation system used by your district to assess their performance and effectiveness. The study will also allow the researcher to examine the extent which feedback from teacher evaluation systems address special education teacher practice and professional development needs.

Why are you being asked to take part?

You are being asked to participate in this research study because you have been identified by Pinellas County Schools as a special education teacher whose primary responsibility is to work with students identified as having disabilities. Your views would provide the appropriate insight into how special education teachers perceive the teacher evaluation system used in your district.

### **Study Procedures**

If you take part in this study, you will be asked to complete an anonymous online survey through Qualtrics through which data will be collected for the purpose of this research. Your completion of the online survey will not in any way be linked to your identity and no personally-identifying information will be collected.

### **Alternatives / Voluntary Participation / Withdrawal**

You have the alternative to choose not to participate in this research study.

You should only take part in this study if you want to volunteer; you are free to participate in this research or withdraw at any time. There will be no penalty or loss of benefits you are entitled to receive if you stop taking part in this study.

### **Benefits and Risks**

We are unsure if you will receive any benefits by taking part in this research study. This research is considered to be minimal risk.

**Compensation**

We will not pay you for the time you volunteer while being in this study.

**Privacy and Confidentiality**

We must keep your study records as confidential as possible. It is possible, although unlikely, that unauthorized individuals could gain access to your responses because you are responding online.

Certain people may need to see your study records. By law, anyone who looks at your records must keep them completely confidential. The only people who will be allowed to see these records are: the Principal Investigator, the advising professor, and the University of South Florida Institutional Review Board (IRB).

- It is possible, although unlikely, that unauthorized individuals could gain access to your responses. Confidentiality will be maintained to the degree permitted by the technology used. No guarantees can be made regarding the interception of data sent via the Internet. However, your participation in this online survey involves risks similar to a person's everyday use of the Internet. If you complete and submit an anonymous survey and later request your data be withdrawn, this may or may not be possible as the researcher may be unable to extract anonymous data from the database.

**Contact Information**

If you have any questions about your rights as a research participant, please contact the USF IRB at (813) 974-5638 or contact by email at [RSCH-IRB@usf.edu](mailto:RSCH-IRB@usf.edu). If you have questions regarding the research, please contact the Principal Investigator at 813-770-6961 or [gbrobbey@mail.usf.edu](mailto:gbrobbey@mail.usf.edu).

We may publish what we learn from this study. If we do, we will not let anyone know your name. We will not publish anything else that would let people know who you are. You can print a copy of this consent form for your records.

I freely give my consent to take part in this study. I understand that by proceeding with this survey I am agreeing to take part in research and I am 18 years of age or older.



### Appendix 3: Initial Invitation to Participants



Pro # 00039485

Dear Colleague,

I am a doctoral candidate in the College of Education at the University of South Florida and a former special education teacher currently completing a dissertation study for a doctoral degree in Special Education Policy Studies.

The purpose of the dissertation is to explore special education teacher perspectives and voices relating to the teacher evaluation system that is used to assess their performance and effectiveness. It is my intention to survey all special education teachers at the secondary level, both middle and high school, in your district. The information obtained from the study may contribute to a change in policy and inform practice and professional development considerations.

You have been identified as a special education teacher with relevant expertise and unique perspectives relating to the teacher evaluation process in your district. Your participation is entirely voluntary. I am not sure you will derive any direct benefits; however, your participation will provide much needed insight into how special education teachers have fared and responded to the evaluation process.

*If you are a former ESE teacher or your **current** primary assignment is **not** teaching students with disabilities, please disregard this survey request.*

Please feel free to contact me at [gbrobbey@mail.usf.edu](mailto:gbrobbey@mail.usf.edu), if you have any questions. Thank you in advance for your help and participation.

Please click the following link to take the 5 – 10 minute survey:  
[https://usf.az1.qualtrics.com/jfe/form/SV\\_a4vc1LKjZV1szOZ](https://usf.az1.qualtrics.com/jfe/form/SV_a4vc1LKjZV1szOZ)

Gordon Brobbey  
Doctoral Candidate  
College of Education  
University of South Florida  
[gbrobbey@mail.usf.edu](mailto:gbrobbey@mail.usf.edu)

#### **Appendix 4: First Reminder Email**

Dear Colleague,

Last week, I sent you an invitation to share your perceptions regarding your district's evaluation rubric that is used to assess your effectiveness as a special education teacher. If you have not yet completed the survey, I would be appreciative if you could take a moment to share your precious insights.

I am a doctoral candidate in the College of Education at the University of South Florida and a former special education teacher currently completing a dissertation study for a doctoral degree in Special Education Policy Studies.

The purpose of the dissertation is to explore special education teacher perspectives and voices relating to the teacher evaluation system that is used to assess their performance and effectiveness. It is my intention to survey all special education teachers at the secondary level, both middle and high school, in your district. The information obtained from the study may contribute to a change in policy and inform practice and professional development considerations.

You have been identified as a special education teacher with relevant expertise and unique perspectives relating to the teacher evaluation process in your district. Your participation is entirely voluntary. You may not derive any direct benefits; however, your participation will provide much needed insight into how special education teachers have fared and responded to the evaluation process.

If you are a former ESE teacher or your current primary assignment is not teaching students with disabilities, please disregard this survey request.

Please feel free to contact me at [gbrobbey@mail.usf.edu](mailto:gbrobbey@mail.usf.edu), if you have any questions. Thank you in advance for your help and participation.

Please click the following link to take the 5 – 10 minute survey:  
[https://usf.az1.qualtrics.com/jfe/form/SV\\_a4vclLKjZV1szOZ](https://usf.az1.qualtrics.com/jfe/form/SV_a4vclLKjZV1szOZ)

Gordon Brobbey  
Doctoral Candidate  
College of Education  
University of South Florida  
[gbrobbey@mail.usf.edu](mailto:gbrobbey@mail.usf.edu)

## Appendix 5: Second Reminder Email

Dear Colleague,

Please consider sharing your thoughts as a contribution to a much-needed conversation about the special education teacher evaluation process in your district. I recognize that you are a very busy person and may not have had the chance to respond to the two previous invitations. I would like to extend one more invitation for you to add your voice to the conversation by completing the survey.

I am a doctoral candidate in the College of Education at the University of South Florida and a former special education teacher currently completing a dissertation study for a doctoral degree in Special Education Policy Studies.

The purpose of the dissertation is to explore special education teacher perspectives and voices relating to the teacher evaluation system that is used to assess their performance and effectiveness. It is my intention to survey all special education teachers at the secondary level, both middle and high school, in your district. The information obtained from the study may contribute to a change in policy and inform practice and professional development considerations.

You have been identified as a special education teacher with relevant expertise and unique perspectives relating to the teacher evaluation process in your district. Your participation is entirely voluntary. You may not derive any direct benefits; however, your participation will provide much needed insight into how special education teachers have fared and responded to the evaluation process.

If you are a former ESE teacher or your current primary assignment is not teaching students with disabilities, please disregard this survey request.

Please feel free to contact me at [gbrobbey@mail.usf.edu](mailto:gbrobbey@mail.usf.edu), if you have any questions. Thank you in advance for your help and participation.

Please click the following link to take the 5 – 10 minute survey: [https://usf.az1.qualtrics.com/jfe/form/SV\\_a4vcILKjZV1szOZ](https://usf.az1.qualtrics.com/jfe/form/SV_a4vcILKjZV1szOZ)

Gordon Brobbey  
Doctoral Candidate  
College of Education  
University of South Florida  
[gbrobbey@mail.usf.edu](mailto:gbrobbey@mail.usf.edu)

---

## Appendix 6: University IRB Approval



RESEARCH INTEGRITY AND COMPLIANCE  
Institutional Review Boards, FWA No. 00001669  
12901 Bruce B. Downs Blvd., MDC035 • Tampa, FL 33612-4799  
(813) 974-5638 • FAX (813) 974-7091

3/7/2019

Gordon Brobbey  
Teaching and Learning  
Tampa, FL 33612

RE: **Exempt Certification**

IRB#: Pro00039485

Title: The Teacher Evaluation Conundrum: Examining the Perspectives of Special Education Teachers

Dear Mr. Brobbey:

On 3/7/2019, the Institutional Review Board (IRB) determined that your research meets criteria for exemption from the federal regulations as outlined by 45 CFR 46.104(d):

(2) Research that only includes interactions involving educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior (including visual or auditory recording) if at least one of the following criteria is met: (i) The information obtained is recorded by the investigator in such a manner that the identity of the human subjects cannot readily be ascertained, directly or through identifiers linked to the subjects; (ii) Any disclosure of the human subjects' responses outside the research would not reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, educational advancement, or reputation; or (iii) The information obtained is recorded by the investigator in such a manner that the identity of the human subjects can readily be ascertained, directly or through identifiers linked to the subjects, and an IRB conducts a limited IRB review to make the determination required by 45 CFR 46.111(a)(7).

As the principal investigator for this study, it is your responsibility to ensure that this research is conducted as outlined in your application and consistent with the ethical principles outlined in the Belmont Report and with USF HRPP policies and procedures.

Please note, as per USF HRPP Policy, once the exempt determination is made, the application is closed in ARC. This does not limit your ability to conduct the research. Any proposed or anticipated change to the study design that was previously declared exempt from IRB oversight must be submitted to the IRB as a new study prior to initiation of the change. However, administrative changes, including changes in research personnel, do not warrant an Amendment or new application.

We appreciate your dedication to the ethical conduct of human subjects research at the University of South Florida and your continued commitment to human research protections. If you have any questions regarding this matter, please call 813-974-5638.

Sincerely,

A handwritten signature in black ink, appearing to read "Kristen Salomon", written over a horizontal line.

Kristen Salomon, Ph.D., Chairperson  
USF Institutional Review Board